SOME ASPECTS OF WATERSHED HYDROLOGY

March 1970 ARS 41-157

Agricultural Research Service

UNITED STATES DEPARTMENT OF AGRICULTURE

"All streams run to the sea,
but the sea is not full;
to the place where the streams flow,
there they flow again.

All things are full of weariness:
a man cannot utter it;
the eye is not satisfied with seeing,
nor the ear filled with hearing.
What has been is what will be,
done;
and there is nothing new under the sun"

Ecclesiastes 1. 7-9

FOREWORD

The Soil and Water Conservation Research Division of the Agricultural Research Service asked the author, Dan Zaslavsky, "to serve as a special consultant in the formulations of theory and premises for initiating a program of research in lateral flows through soil mantles and through deep lying aquifers." This publication reports part of the results of this assignment (August 1967-September 1968). It brings together a series of concepts and ideas concerning the management of water and soil in the context of agricultural watersheds. The subject is complex and involves many aspects and disciplines.

Longtime workers in a field sometimes lose sight of the broad picture. Dr. Zaslavsky was an outsider, not directly involved previously with the field of watershed hydrology. The observations here recorded are the result of 6 months of intensive review of the literature and of ongoing research, and of extensive conversations with specialists in the field. They cut across the conventional bounds of hydrology, and touch on erosion, soil science, and principles of experimentation.

J. van Schilfgaarde
Associate Director

FTS 323 - 5265

CONTENTS

Page

LIST OF FIGURES

Page

SOME ASPECTS OF WATERSHED HYDROLOGY

By Dan Zaslavsky[1]

Chapter 1 THE CONCEPT OF WATERSHED ENGINEERING

1.1 The Need for Well-Defined Terms

A term is meant to give a descriptive notion of a system, a process, or both. Certainly, colloquial use of loosely defined terms adds to the coloration of the language, but in scientific thought, new terms are used for more economical storage, retrieval, and delegation of information. Unless these terms are clearly and rigorously defined they can confuse, distract, or delude those receiving or giving the information. Loosely defined terms are especially harmful to the beginning student or to scientists in associated fields.

To justify this discussion, one may refer to the recently published Water Resources Thesaurus (58)[2] for some examples of linguistic confusion. For many entries, the thesaurus indicates a prefix "used for" or "use." These are mostly cases where, in the authors' opinion, the terms have been misused.

An example is "adhesion" and "cohesion." Adhesion and cohesion are the same, physically, except for the point of view or scale of observation. What is adhesion microscopically, is cohesion macroscopically. The same is probably true for adsorption and absorption. Now, this scale distinction leads us directly to the problem of defining "aggregates." About the best we can do is to say that an aggregate is a piece of soil that has a cohesion larger than adhesion. The relative sense of the words leads to an oversimplified, almost primitive concept of soil structure. (See chapter 7.)

Another term in the Water Resources Thesaurus is "agricultural watershed." What has agricultural to do with it? Is it an administrative definition that is curtailing our horizons?

The most abused word is probably "average." There are numerous types of averages. The physical situation and the purpose of the analysis determine the type of average to be used. (See chapters 3 and 4.) Very rarely, in papers on hydrology, can one find a proper definition and justification of the average. This is probably the one largest source of errors and artificial production of themes concerning misinterpreted experimental results. A large part of this report will be devoted to this term. (See chapters 3, 4, and 5.)

The terms "runoff," "subsurface runoff," and "interflow" are especially vague. The thesaurus states ... "Use surface runoff for diffused surface water" ... "Use subsurface runoff for interflow," and so forth. As yet there is not really a dependable physical definition of these. The terms "interflow" or "quick return flow" imply that water gets into the ground and then back out. In fact, all computations separating runoff and interflow are deduced from

---

[1] Special consultant, Soil and Water Conservation Research Division, Agricultural Research Service, U.S. Department of Agriculture.

[2] Underscored numbers in parentheses refer to References Cited at end of publication.

flow measurements in a water channel by some arbitrary and artificial model and curve fitting without physical significance. It is questionable whether such a separation is at all possible or necessary. The separation between terms and the proper delay times are changed at least by the size of the area studied. Originally, the term "interflow" was (probably) used by some, only as a notation for an intermediate component in some arbitrary breakdown of an experimental hydrograph. The authors of the thesaurus (57) suggested calling it subsurface runoff instead. Thus, they not only produce a self-contradictory type of word, but formally establish a misinterpretation. Various authors, attempting a physical decoration interpretation of their otherwise worthwhile system analysis, have produced figurative diagrams of the routes taken by the water. By doing this, they have invited both an undue criticism of their cubistic embodiment of flow processes and an ever-complicating almost surrealistic concept of nature.

The basic difficulty in the hydrologic terms given above is that they are intuitive. A large part of this report is devoted to a more systematic way of defining a hydrologic model by applying mathematical procedures to the basic physical phenomena.

1.2 Definition of a Watershed

Most generally, a watershed can be defined as a body of soil with definite boundaries around it, above it, and below it. A <u>positive water accretion</u> to its upper boundary is in the form of precipitation and a <u>negative accretion</u> is in the form of evaporation. There can be <u>drainage</u> laterally or vertically when water runs out of it and <u>leakage</u> when there is a perched water horizon.

There is little interest in defining just any body of soil. The name "watershed" implies that it should be defined in such a way as to enable formulation of a water regime with emphasis on precipitation and water yield.

Any flow problem must be formulated according to certain rules. It must include the exact definitions of boundaries and the specification of the boundary conditions. Some boundaries would be more convenient to have than others. For example, on an arbitrary boundary surface, one has to specify the interaction between the two bodies on the two sides of the boundary. The procedure is often difficult or impossible. On the other hand, the boundaries may be chosen so that there will be little or no interaction and at most an interaction that can be easily approximated. The definition of a <u>drainage basin</u> attempts to meet these requirements.

1.3 Definition of a Drainage Basin

Generally <u>watershed</u> and <u>drainage basin</u> are used synonymously. In British literature "catchment" is used for a drainage area (<u>60</u>, <u>sec. 14-2</u>). However, watershed seems to be a more loosely defined concept, especially in the light of engineering application.

From a hydraulic point of view, a drainage basin can be defined as the space (usually soil) from which water flows into a certain <u>outlet</u> or <u>collector</u>. The collectors may be in a form of points, lines or surfaces that are crossed by streamlines.

A surface through streamlines or a <u>stream surface</u> has an outstanding property in that there is no water flow across it. Thus, one can cut or partition a soil body with flowing water in two parts at a stream surface without influencing the flow pattern in either part.

Starting at a collector and following along a streamline, there will be a monotonic rise in the flow potential. At some point the streamline will reach a maximum potential or a stagnant point. It is convenient to limit the drainage basin at the stagnation points since the flow there is specified--it is zero. In hydrology the geometric locus of these stagnation points is called the divide line.

Another type of boundary is a surface where the influx or outflux can either be measured or estimated. In the upper surface, precipitation and evaporation must be determined. At the lower surface, leakage or drainage has to be known (often they are part of the flow problem to be solved). Some feeding veins, springs, and such, have to be measured separately.

In summary, a drainage basin is defined first by a certain collector. Most of the boundaries are then defined as streamlines that stretch between this collector and a divide line, or a source. The upper and lower surfaces are mostly considered as stream surfaces with a superposition of known or estimated accretion.

1.4 A Topographic or Surface Watershed

Some practical procedure has to be found to delineate the drainage basin in the field. The most common way is the topographic or surface delineation. It is in this sense that almost all hydrologists define a watershed.

The procedure is as follows. A collecting point or collecting line is chosen on a topographic map. From the collector one draws lines normal to the elevation contour lines until they reach a topographic divide line.

If the soil surface is considered to be impermeable (in a normal direction) and isotropic (parallel to the surface), the lines normal to the contour lines can be considered as streamlines. They define then topographic flow tubes. In this case the flow system is completely defined by (a) the collector, (b) surface streamlines, and (c) a divide line.

By moving the collector downstream, the area of the surface watershed is increased in two possible ways.

a. The boundary streamlines get farther apart and the divide line becomes longer. It can be said that the number of flow tubes increases. There are singular points and singular lines where different streamlines converge. These convergence lines are usually at a minimum elevation or at a channel's route. The collectors are commonly placed on such lines. Moving downstream on the minimum line, the catchment area will increase gradually.

b. At some points the minimum lines or singular lines branch off. At these points there is a marked increase in watershed area due to the additional stream tubes. It is usually preferred to define a watershed by placing the collector just above a minimum line or a singular line of water convergence.

1.5 Some Consequences and Limitations

The watershed as commonly defined is strictly of a topographic nature. Very often it has very little to do with the actual flow net, the estimated water volumes, or the rates of discharge. Inconformity between the geological layers and the soil surface can cause interception of water. Springs may contribute to surface flow. Leakage may be unaccounted for. Anisotropic conditions will cause streamlines to deviate from the topographic streamlines.

The incompatibility between the topographically defined watershed and the actual flow configuration increases as watershed size decreases. It is always possible to combine small watersheds to form a larger one by moving the outlet or collector point downstream. The ultimate downstream movement is at a point or line of discharge into the ocean. The inconformity is expected then to be the smallest, especially if the collector is a long line. The definition of a surface watershed is closely related to the artificial distinction between surface runoff and subsurface runoff.

The soil surface is often anisotropic. By this, one means that the actual streamlines will not be parallel to the surface slope (or elevation gradient). This will occur wherever there are terraces on the slope, furrows, transitions from one soil type to another, naturally occurring gullies or stream beds that are not normal to the contour lines. In these cases, even a flow above the surface will not be normal to the contour lines, and the definition of surface watershed loses its merit. The discrepancy caused by surface anisotropy is expected to be the largest on the small watersheds. The error in estimating the total water volume is expected to be smaller on a larger watershed.

Agricultural watersheds are considered for some reason to be small watersheds defined by the surface topography. This makes them the most incompatible type from a hydrological point of view. Still, there may be practical reasons to support such a unique concept.

1.6 Watershed Engineering

The advantage in defining a drainage basin is that solutions of water flow problems are not affected if impermeable boundaries are assumed to coincide with streamlines. However, other flow processes such as dispersion of ions, heat exchange, and electrical currents may follow an altogether different flow net. Thus, behind the definition of a drainage basin there is the notion that convective terms due to water flow are major factors in all other processes. This is most often justified. However, care must be taken when other flow processes become important.

The actual definition of a watershed usually assumes a certain physical model and is thus even more limited than a drainage basin in its applicability. In the above, the topographically defined watershed was shown to be applicable only for surface flow on an isotropic soil surface.

The coined phrase "watershed engineering" opens up any number of possible definitions based on other physical models, such as storm areas, pollution areas (from runoff or waste disposal), ground water drainage areas, or geological configurations.

In addition, definitions may be obtained through operational considerations, such as administrative boundaries, translocation of water from wet to dry areas, larger scale handling of resources concerning water quality and economy, agricultural practices and potentials, and industrial and domestic needs.

It must be understood that a canal, a dam, a road, a pipeline, and a simple pump are just as natural today as naturally occurring surface slopes but they negate completely the usual definition of a watershed. We may use the watershed definition as a naturally occurring drainage basin only if we are sure that it does not produce a bias in our education, research framework, and engineering practices. One should never forget that part of the practical problem is to define the artificially induced interactions between naturally occurring drainage basins, and that even a drainage basin is rarely a closed system.

Chapter 2 THE COMPLEXITY OF THE SYSTEM INVOLVING HYDROLOGY

2.1 The Scope

Following water in its natural cycle and in its uses, one can hardly avoid any field in human endeavor. In the study of the role of water, three main levels of effort can be recognized:

1. Observation and descriptive organization of data
2. Physical interpretation and formulation
3. Applications or intervention by man to change conditions for his benefit

The descriptive phase becomes more and more voluminous. Compilation of data competes (successfully?) with our fast-increasing capacity to handle them. There is very little fundamental about the study of water today. Most of our physical formulations must make some sacrifice in accuracy to be able to cope with the complexity of the media and the inputs. Therefore, data collection and physical studies are aimed mainly at certain specific applications or modifications by engineering.

The range of applications of possible decisions and their consequences is wide. If one faithfully followed the proliferation of interacting fields of interest, he would before long cover every corner of human life on earth. This is certainly not practical.

2.2 Curtailing a System into a Subsystem

The problem facing the hydrologist is how to curtail his system to obtain a balance between several contradicting demands. The formulation should be wide in scope but simple enough for use. While attempting to obtain a comprehensive account, one should give attention to the need for reaching decisions in limited time and by feasible effort.

In view of the three levels mentioned in section 2.1, the observational, the physical, and the practical, there are also three ways to curtail the system: (1) data collection, (2) physical reasoning, and (3) practical considerations.

2.3 Curtailing by Data Collection

One obvious way to curtail the system is to decide arbitrarily what information to collect on the basis of familiar disciplines or available measuring equipment. Usually this system of curtailing yields data of a particular classification and is arbitrary because the whole physical system is not elucidated. It is thus quite common to have a long detailed record of rainfall with no data for runoff; or to have data of temperature, humidity, and wind, but no soil moisture data, or physical characteristics that would enable a physical analysis or even a vague corroboration of the evaporation estimates. As a consequence, the physical analysis and the decisions made are limited in scope, or worse, biased. Unfortunately a more or less arbitrary and limited range of data will serve as a template for our subsequent analysis and not the other way around.

Rain measurements seem to be simple; this is why we have so much rain data, much of which is of limited use. However, accepting that the making of measurements should be continued, definite improvements in measurement and handling of data are needed.

2.4 Physical Reasoning in Curtailing the System

A better way to reduce our system to smaller subsystems is through some physical considerations. There are several dimensions to the system: actual time and space coordinates, and many processes that take place simultaneously. One can simply subdivide the system into subsystems of smaller dimensions in time and space if there is a negligible interaction between subsystems (26), or if, when the interaction is known, it can be estimated or assumed. In other words, one must have boundary and initial conditions to be able to find solutions. (See, for example, sections 1.3, 1.4, 9.2-9.5.) The hydrological problem must be "well posed" (4, section 6.7). Often, when the subsystem is large enough, certain errors in defining the boundary conditions make little difference as far as the results in the bulk are concerned. Time is of special interest here, because many processes have a diminishing influence with time or depend very little on past history.

As for the complexity of the physical picture it is always necessary to neglect certain processes. However, those neglected should first be proved to have little influence on the parameters of interest. Often our guess as to such influences is not rigorously demonstrated. (See sections 3.8, 3.9, and the end of 3.7.)

One must not attach too much fundamental meaning to any one hydraulic model. Rather, the model has to prove useful with respect to some practical need, and it must be shown to be rigorously related to a real physical picture. Some degree of trial and error is justified in providing approximate solutions to a somewhat idealized, abstract model and later to search for the possible cases that fit the solution. However, there is too much "substitution of glass beads for soil," which is a poor approximation for anything except for glass beads themselves. It is worth remembering that flow in the field is not one-dimensional, not linear, not uniform, not isotropic, not steady, and not saturated. The difficulty in solving a problem is not by itself a sufficient justification for oversimplified laboratory experiment.

Many attempts are being made to study the processes in a small space element and at a given moment. Where would we be today if we tried to study the microscopic behavior of interfaces before even attempting to drill a well? The "rational" formula for runoff where the runoff is proportional to the rain but the coefficient of proportionality changes as a function of time, watershed size and rain intensity seems to be an example of a useful macroscopic law without microscopic physical interpretation.

2.5 Curtailing by Practical Considerations

The practical applications of a research study are often neglected altogether by soil physicists and hydrologists. Still, a practical objective often serves as the simplest and fastest way to reduce the complexity of the system. Erosion is an excellent example. There are several ways in which erosion can be reduced. It is often easiest to find which of these best apply to a given case.

It will be shown (chapter 7) that seepage forces, runoff, and raindrop splashing are often concurrent. Moreover, similar means maybe applied to prevent them from eroding the soil. One can choose means to prevent erosion merely by understanding the mechanisms and trying various implements, despite the fact that a quantitative account of the whole phenomenon is not possible at present.

Numerous other problems of drainage, water storage and recharge, sewage and irrigation could be cited where the major decisions can be made and are being made on the basis of limited information. Certain economical, social, or political issues often appear to be determining factors in making a decision. There is also a trial and error in engineering that is often cheaper than research. Too many research workers indulge in "paper-directed" research and too few in followup and analysis of "experimental" engineering. The latter is often the cheaper way to improve the prediction and save on cost. A thorough understanding of the most intricate physical phenomena with a sense for practical application can produce most fruitful results in the shortest time even without a comprehensive quantitative formulation. The intermediate approach is most laborious and sterile. It lacks physical rigor because it is too complicated and avoids practicality on the pretense of fundamentality or comprehensiveness.

2.6 The Stochastic Progress in Research

The definition of a subsystem by a given researcher or engineer is too often a result of personal limitations, likes or dislikes, with very little objective reasoning. Through some stochastic process there is a spontaneous flow of progress at a given rate. Many results, thus obtained, fail to meet the needs. Taking the changing circumstances as boundary conditions, it takes too long before they manifest themselves in the stochastic progress of research.

The scientific establishment follows to a great extent traditional divisions that originate from political, geographical, and ethnical sources. A system with better internal feedback should be based on the media concerned, with a secondary accent equally divided between the producing organization and the clientele. The important goals in handling soil and water are not identified or stressed enough in many research efforts.

2.7 A Breakthrough as a Simplifying Device

We often do not realize that we work on marginal problems. One of the values in assigning priorities within a program is that it requires the anticipation of the outcome from both the producer's and the clientele's point of view.

An actual breakthrough means that by application of new technology many factors in the system become secondary as far as decision making is concerned. Many of the subjects on which we spend large sums of money today would lose their importance by a breakthrough. One can convince himself in that matter by looking at any technical book that is a few decades old.

Small watershed studies seem to pose a logical contradiction today. On the one extreme, there is a need for studying soil-water interaction in principle and applying new techniques such as in irrigation, filtration, purification, stabilization, and fertility. On the other, one has to consider large systems. A

small watershed is often not a pertinent unit from either point of view. Worldwide balances become more and more important. Many new factors in human health and welfare are just as important today as terraces and gravity irrigation were in the past.

The development of some unconventional water sources may make certain aspects unimportant in water resources as we know them today. In the past, conservation measures were dictated by the need to maintain soil fertility. Today they are dictated more by city people who wish to maintain their recreational activities and drink good water (at least in the United States). This is due to breakthroughs in agricultural practices that occurred in the last 30 years and have nothing to do directly with water in general or with hydrology specifically.

Priorities must be repeatedly reexamined. The training of soil and water scientists must be less descriptive and more physical and mathematical so that they will be able to perform better in a given subject and move more easily to another one. Such a change in curriculum seems to be the best hope for persuading capable students to enter this field.

The techniques of operations research are certainly important. They should be applied to our system's analysis. However, while they offer useful concepts and techniques that should guide every engineer, they can also serve as a refuge for those who find the physics or the engineering difficult or less glamorous. Recent curricula of water resources often produce engineers that are not capable of solving a flow problem but can talk about it. They could streamline the operation of a series of reservoirs if someone would be there to operate them.

Chapter 3 A STATISTICAL TREATMENT OF THE WATERSHED PARAMETERS

3.1 Introduction

Practice shows that the various components and parameters involved in the hydrologic cycle usually have a range of values varying in time and space. This may introduce an error or a bias in our estimates depending on how we obtain average values for these parameters. This problem is quite general and may be stated as follows: If a certain output Y is a function of a number of inputs Q_i ,

$$Y = Y\,(Q_1, Q_2, \ldots Q_i, \ldots Q_n), \tag{3.1}$$

then with few exceptions

$$\bar{Y} \neq Y\,(\bar{Q}_1, \bar{Q}_2 \ldots \bar{Q}_i \ldots \bar{Q}_n), \tag{3.2}$$

where $\bar{Y}$ and $\bar{Q}_i$ indicate, respectively, an average Y and average parameters Q_i over time, space, or both.

One exception occurs when 3.1 is a linear and homogeneous equation and there are no bounds on the domain of Q_i such as

$$aQ_i + bQ_j > \text{Constant.} \tag{3.3}$$

The right-hand side (r.h.s.) of equation 3.2 can at most be an approximation of the left-hand side. The connection between the two sides must be found in terms of the distribution of Q_i or in terms of their variance, covariance, and derivatives of different order of Y with respect to Q_i. Thus, quite generally, the parameters of our system are not only $\bar{Q}_i$ but also the modes of Q_i distribution. Unless proved negligible, they must be considered equally important.

In the following, we shall try to demonstrate how these distributions can be accounted for and how possible errors can be estimated a priori. For a hydrologist and a soil scientist, the notion of inhomogeneity and an average entity should always be kept in mind. It should become a standard practice to check their influence.

As a consequence of such analysis, one may have to choose the type of required measurement or instrumentation, the size of a sample, or the number of samples needed to obtain a representative value. The analysis may show a certain procedure to be impractical and lead to the conclusion that one cannot improve our statistical sampling and have to allow certain inherent error or randomness in our study. Finally, as will be indicated throughout this report, being aware of such inhomogeneity may lead to a better physical understanding of the system's nature. Even if one cannot get a quantitative account of the inhomogeneity as a routine engineering practice, he can at least attempt to avoid serious mistakes.

A certain output $Y(Q_i)$ can be a linear function of inputs Q_i for any given point in time or space. However, the average, $\bar{Y}$ can become a nonlinear or nonhomogeneous function of average inputs $\bar{Q}_i$. Thus, when we attempt to relate $\bar{Y}$ to $\bar{Q}$, we might have to deal with a nonlinear function.

Too often an average is taken to be synonymous with arithmetic average. This, of course, is the simplest one and often serves as a good approximation.

Generally, rain measurements, evaporation, slope, and infiltration are used in terms of their arithmetic averages with no weight function. As will be shown, this may lead to great errors.

3.2 The Fluctuation Response Index

The concept of fluctuation response index (F.R.I.) was introduced by Zaslavsky and Buras (69) and Zaslavsky and Mokady (71). The index describes quantitatively the sensitivity of the mean to fluctuations and was used to evaluate the crop yield response to nonuniform application of irrigation water and fertilizers.

Consider equation 3.1 and express each Q_i as the sum of the averages $\bar{Q}_i$ and deviations from the average, q_i. Expanding Y around $\bar{Q}_i$ in Taylor's series we obtain

$$Y(\bar{Q}_i + q_i) = Y(\bar{Q}_i) + \Sigma \frac{\partial Y}{\partial Q_i} q_i + \frac{1}{2!} \sum_{ij} \frac{\partial^2 Y}{\partial Q_i \, \partial Q_j} q_i q_j$$

$$+ \frac{1}{3!} \sum_{ijk} \frac{\partial^3 Y}{\partial Q_i \, \partial Q_j \, \partial Q_k} q_i q_j q_k + \ldots \tag{3.4}$$

Averaging over an area A (integration and division by A) one notes that $Y(\bar{Q}_i)$ and the derivatives of Y are fixed values for a given set of $\bar{Q}_i$ values. By definition the second term on the right-hand side of equation 3.4 vanishes on averaging. The fourth term would also vanish if Q_i distribution were symmetrical around $\bar{Q}_i$. On averaging, the third term in equation 3.4 becomes

$$\frac{1}{2!} \sum_{\substack{i \\ i=j}} \frac{\partial^2 Y}{\partial Q_i^2} \frac{\int q_i^2 \, dA}{A} + \frac{1}{2!} \sum_{\substack{ij \\ i \neq j}} \frac{\partial^2 Y}{\partial Q_i \partial Q_j} \frac{\int q_i \, q_j \, dA}{A}. \tag{3.5}$$

Obviously, the first term in equation 3.5 is the product of the second derivative and the variance σ_i^2 of the Q_i. We call

$$\frac{1}{2!} \frac{\partial^2 Y}{\partial Q_i^2} \qquad \text{the response index (RI)} \tag{3.6}$$

and

$$\frac{\int q_i^2 \, dA}{A} = \sigma_i^2 \qquad \text{the fluctuation indes (FI).} \tag{3.7}$$

The whole term is then the fluctuation response index. The second term in equation 3.5 gives the covariance of Q_i and Q_j times the mixed second derivative. The covariance may be written as $\sigma_{ij} = \sigma_i \sigma_j \rho_{ij}, i \neq j$, where ρ_{ij} is the correlation coefficient (23, p. 540). Thus:

$$\bar{Y} = Y(\bar{Q}_i) + \frac{1}{2!} \Sigma \frac{\partial^2 Y}{\partial Q_i^2} \sigma_i^2 + \frac{1}{2!} \sum_{\substack{ij \\ i \neq j}} \frac{\partial^2 Y}{\partial Q_i \partial Q_j} \rho_{ij} \sigma_i \sigma_j. \tag{3.8}$$

There is some difficulty in evaluating the derivatives of Y. In the field, we often obtain the values of average output $\overline{Y}$ over an area. It can be shown, however, that the derivatives of $\overline{Y}$ often will not be much different than the derivatives of Y. These derivatives are used in the correction terms that are often a fraction of $\overline{Y}$. Thus, the use of derivatives of $\overline{Y}$ will involve an error which is only a fraction of a fraction and is expected to be negligible. It may also be shown that the difference between $\partial^2 Y/\partial Q^2$ and $\partial \overline{Y}^2/\partial Q^2$ is largest when higher derivatives of Y are not negligible. When this happens 3.8 is insufficient and additional terms are needed.

Before we cite examples (sections 3.3 to 3.6, 4.4), let us consider three more pertinent points. First, equation 3.8 can be expressed in a dimensionless form by dividing it through by $Y(\overline{Q})$ and by dividing and multiplying the second and third terms by $Q_i Q_j$. Thus, one obtains in general

$$\overline{y} = y(\bar{\theta}_i) + \frac{1}{2!} \sum_{\substack{i \\ i=j}} \frac{\partial^2 y}{\partial \theta_i{}^2} \sigma^2_{\theta i} + \frac{1}{2!} \sum_{\substack{ij \\ i \neq j}} \frac{\partial^2 y}{\partial \theta_i \ \partial \theta_j} \rho_{ij} \quad \sigma_{\theta i} \ \sigma_{\theta j} \tag{3.9}$$

where y is the relative output, θ a relative input, $\sigma^2_{\theta i}$ the relative variance of θ or the coefficient of variation squared, and ρ_{ij}, the correlation coefficient between θ_i and θ_j.

$$\theta_i = \frac{Q_i}{\overline{Q}_i}, \tag{3.10a}$$

$$y = y(\theta_i) = \frac{Y}{Y(\overline{Q}_i)}, \tag{3.10b}$$

$$\sigma_{2\,\theta i} = \frac{\int (\Delta Q_i)^2 \, dA}{\overline{Q}_i{}^2 \int dA} \tag{3.10c}$$

In the above, $\overline{Q}_i$ was taken as an instantaneous average input over the whole area. The same calculation can be done for an average over the area and time where $\overline{Q}_i$ becomes a time and space average input. Q_i may thus be averaged over a limited time and space:

$$\overline{Q}_i = \int_{t_1}^{t_2} \int_{y_1}^{y_2} \int_{x_1}^{x_2} Q_i \ dx,\ dy,\ dt. \tag{3.11}$$

This is a subsystem average which smoothes out part of the fluctuations (those with higher frequency). Any changes of $\overline{Q}$ in time and space will be considered relevant to our problem for which we seek a detailed characterization. Thus, the range of averaging may be different for different purposes. This subsystem averaging is useful in defining a continuum (section 5.3).

At times, temporal fluctuations will be smoothed out by a spatial integration. This would then be some ergodic random process and obeying some ergodic hypothesis. Whether it will be so regarded will depend on the autocorrelation of time fluctuations at increasing distances. The correlation between time fluctuations at two points can be estimated with an increasing distance and with a different time phase between them. This is valid until the correlation diminishes. If the size of a subsystem is larger than this distance, then the time fluctuation will be smoothed out by a space averaging.

3.3 Local Averaging Capacity

The fluctuations of an input property Q_i may have a range of frequencies, each with random amplitudes. Each frequency contributes to the variance of the quantity Q_i. The variances are additive. This is very much like the power distribution of noise where the electrical potential of one fluctuation is V and the power involved is proportional to V^2.

The system itself has a capacity to smooth out some of the fluctuations. For example, consider rain excess accumulated in small puddles (Q_i here may be the rain, the infiltration or the surface storage). These puddles can occur long before the average rain intensity exceeds the average infiltration capacity. This is because of local fluctuations. However, in the field, some puddles will have an outlet and some will not. Those that have an outlet may empty into spots of higher infiltration capacity. Thus, some of the local fluctuation will be smoothed out. In addition, there will be some correlation between the parameters. This problem will be similar to electronic noise transmitted through a rectifier of a certain frequency range. Thus, in our analysis of the F.R.I., we should use the effective variance σ_{ei}^2 which is smaller than σ_i^2. It is expected that the higher frequencies are damped more easily. If the size of the subsystem allows measurement of fluctuations of a given frequency range, there is no way of predicting with any degree of accuracy about larger systems or other frequencies. As an example, if we take daily rain measurements over a year, we cannot infer about minute by minute fluctuations or about climatic cycles that take a decade or two.

3.4 The Use of F.R.I. in Dimensionless Form in an Infiltration Equation

Examples concerning irrigation are given by Zaslavsky and Buras (69) and work dealing with different aspects of chemical fertilizers, especially phosphates, is described by Zaslavsky and Mokady (71).

We shall take here as one example the infiltration formula used by the USDA Hydrograph Laboratory (15)

$$(f - f_c) = aF^n, \tag{3.12}$$

where $(f - f_c)$ is the net recharge of topsoil or inflow minus outflow. F is the total available storage space for water at a particular time and a and n are experimental parameters. Taking

$$\frac{\partial^2 (f - f_c)}{\partial F^2} = an(n - 1)F^{n-2}, \tag{3.13}$$

we may write from equation 3.9 and 3.10

$$\bar{y} = \frac{(f - f_c)}{(f - f_c)_{\bar{F}}} = 1 + \frac{1}{2!} \frac{an(n - 1)\bar{F}^{n-2} F^2}{a\bar{F}^n} \frac{\int (\Delta F)^2 dA}{\bar{F}^2 \int dA} \tag{3.14}$$

$$\bar{y} = 1 + \frac{n(n - 1)}{2} \sigma^2_{F/\bar{F}}. \tag{3.15}$$

Here $\bar{F}$ is the average storage capacity in the soil, $\sigma_{F/\bar{F}}$ is the relative variance, and $\bar{y}$ the relative average rate of storage decrease over the field. Holtan (15) experimentally found that n = 1.4. Thus,

$$\bar{y} = 1 + 0.28 \quad \sigma^2_{F/\bar{F}}. \tag{3.16}$$

F is not the total aerated porosity in the profile (per unit area of the field), but rather a function of a moisture content at a particular location. Because of differences in antecedent moisture and soil properties over an area, the dimensionless $\sigma_{F/\bar{F}}$ can be even as high as 0.5. We may thus anticipate a bias in our initial estimate of $f - f_c$ of as much as 14 percent. Equation 3.12 is useful only when F can be measured or estimated, and thus the rate of infiltration predicted. What was proved in equation 3.16 is that if $\bar{F}$ would be available from direct measurement or estimates, a large error can occur unless fluctuations are somehow taken into account. Once equation 3.12 is integrated the cumulative error may become larger.

By using a measured arithmetic average value of initial storage capacity F_o in Holtan's (15) formula, we may make a systematic error in calculating F. This error would be largest for soils of high drainable porosity.

In equation 3.12, one has to measure not only $\bar{F}_o$ but f and f_c at t = 0 and consider f_c to be also some function of F. Thus, more generally at a point in the field

$$f_i = a_i F_i^{\,n_i} + fc_i\,(F, fco_i, p_i) \tag{3.17}$$

where i indicates a point in time and space, fco is initial f_c and p is the precipitation. Variations in a_i, n_i, F_o, and p will cause the average infiltration $\bar{f}$ to vary from the one obtained from average a, n, F_o, f_c, and p. The effort to produce a formula of this form (equation 3.12) is justified if, in the future, one will be able to assign values F_o, Fco, and a to the field, from independent measurements and then correlate the rain with the infiltration, storage, and runoff. In this case, one will also have to include the F.R.I. analysis for a multivariable case where a number of factors will contribute to the bias. Accuracy in the infiltration equation is required because of its use in runoff predictions. Basically, the runoff is obtained by subtracting the infiltration from the rain and integrating with time. The total infiltration most often is 80 to 90 percent of the rain and more. Thus, an error of 10 percent in the estimate of infiltration constitutes an error of 100 percent or more in the estimate of runoff. Let us assume that a formula such as equation 3.17 is a valid one for a given point in the field at a given time. If this formula is to be used for estimating runoff, two practical conclusions are unavoidable.

1. The accuracy of our measurements of rain and soil parameters must be within a certain fraction ε of the runoff. The runoff makes a certain fraction c of the rain (and most often about the same fraction of the infiltration storage and evaporation combined). The accuracy of our measurement must then be εc. If ε is 10 percent and c is 5 percent, then εc is 0.5 percent. Is this practical? Will it be practical in the future?
2. If one shoots at an accuracy of 1 percent or less, it is mandatory that he will measure not only the averages. At least the second moment has to be measured in the field.

There is little argument about lack of physical significance to equations such as equation 3.12 or 3.17. The question is rather, what is the significance of the tremendous effort expended in fitting the hydrograph curve by arbitrary adjustments of different coefficients?

3.5 The Soil Conservation Service Formula

We use U.S. Soil Conservation Service's (57) formula for runoff as another example. It is

$$Q = \frac{(P - 0.2d)^2}{(P + 0.8d)} \qquad (3.18)$$

where Q is the total runoff, P, the total rain, and d, the storage. Let us estimate the error in Q as a result of fluctuations in P. From equations 3.9 and 3.10a, b, c.

$$\frac{1}{2!} \frac{\partial^2 Q}{\partial P^2} \frac{\bar{P}^2}{\bar{Q}} = \frac{d^2 P^2}{[(P - 0.2d)(P + 0.8d)]^2}. \qquad (3.19)$$

To work out an example, let us assume P = d = 2.0 cm. The relative fluctuation response index from equations 3.19 and 3.9 is then $16/33.1\sigma^2_{P/P_0}$ or about $0.48\sigma^2_{P/\bar{P}}$. Fluctuations in P can be as high as 100 percent and more so that $\sigma^2_{P/\bar{P}}$ can be as high as 0.5 or larger. Thus, by using average values of P in the above formula, we may make a systematic error of 25 percent and more in calculating Q. If fluctuations in d are included, the error can be larger.

An interesting generalization may be discussed at this point. A positive second derivative indicates a concave line of the function Q(P). A negative derivative would indicate a convex line. One can state generally that fluctuations will make the average larger than the function Q(P) when the line is concave; but they would have reduced the average if the line were convex.

Clearly, from equation 3.18 there is a threshold value of P at P = 0.2d. The second derivative in relative terms approaches infinity. This means that at small rains (small P) the bias due to averaging would be the largest. Neglecting the second moment of the rain, one would predict runoff values that are too small. This is really the experience in several predictive models.

If one is interested in very large storms only, then the error involved would be greatly reduced. For example, if P = 4d, the bias would be $16\sigma^2/(1824)^2 \sim 0.048\sigma^2$. It is well worth indicating here that the value of a certain approach may be limited to certain engineering purposes. Someone interested in protective drainage structures and flood control may find the use of arithmetic averages quite satisfying. This is because he is interested only in extreme storms where d is small and P is large. A person interested in water yield may find that small storms add up to a large proportion of the total water volume. Alternatively, one may find that it is not economical to design the storage, transmission, and water treating installations for extreme peaks. Again, he will be interested in smaller storms. Then the rain variance and storage variance should be as important as their averages.

Fluctuations in storage d can also be included in the calculation and will further contribute to the possible bias in estimating the runoff. Neglecting the correlation between P and d equation 3.19 applies also to the d variation.

3.6 Changes in Unsaturated Hydraulic Conductivity

In the future, hydrologists expect to be able to predict both infiltration and evaporation by using some physical properties of flow in unsaturated soil. Following is an example of a possible analysis where the unsaturated hydraulic conductivity can be predicted on the basis of measurements of field capacity and saturated hydraulic conductivity. We take the formula for the hydraulic conductivity K in unsaturated soil (4, ch. 9)

$$K = K_r \left[\frac{C - C_o}{C_r - C_o}\right]^3 \tag{3.20}$$

where K_r is some reference hydraulic conductivity (possibly saturated K). C_o may be taken as the moisture content at field capacity. We define $C' = C - C_o$ and $C'_r = C_r - C_o$, then

$$K_R = \frac{K}{K_r} = \frac{C'^3}{C'^3_r} \tag{3.21}$$

and

$$\text{F.R.I.} = \frac{1}{2} \frac{\partial^2 K_R}{\partial C'^2} \frac{\bar{C}'^2}{\bar{K}_R} \sigma^2_{C'/\bar{C}} = 3\sigma^2_{C'/\bar{C}} \tag{3.22}$$

The F.R.I. depends on the distribution of moisture content, field capacity, and saturated hydraulic conductivity. In a certain field, the antecedent moisture can vary widely. $\sigma^2_{C'_R/\bar{C}_R} = 0.25$ will not be uncommon. Thus, the systematic error made by using equation 3.20 with arithmetic averages could be as high as 0.75 $\bar{K}$.

If one expects to actually measure K distribution or infiltration distribution in the field (6) and derive on average (10), he will definitely make errors of the above order. As Crawford and Lindslay (10) attempted to make predictions with smaller errors, they had to utilize some experimental coefficients to obtain a curve fitting. As the bias is expected to be different for different storms (sections 3.4 and 3.5), there is little hope to get one such coefficient for all storms in the same watershed. Moreover, one can conclude quite generally that assuming the use of arithmetic averages and an experimental correlative model, the experimental coefficients will vary with antecedent moisture, rain intensity, and total rain. Thus, one can also conclude that a linear model of this type is impossible.

It is not surprising that models that defy physical sense often appear successful. One only needs to assume some bell shaped hydrographs or to obtain them from some procedure which places different amounts of rain at different parts of the watershed with the crudest assumption about runoff delay. If the peaks, volumes and skewness are correlated with the rain peak value and skewness and if the base flow is measured or estimated, the chances not to succeed in curve fitting are very small.

3.7 A Statistical Calculation of the Expected Rain Excess (A Linear Model)

In the sections 3.1 to 3.6, I quoted cases that have some type of a non-linearity in their physical behavior. In the following, I shall consider a model which is basically linear but not homogeneous. By this, I mean that the function has a threshold or that the function line does not run through the origin. Even in this case, neglecting the fluctuations around the average would involve a systematic error or a bias.

Consider the simple model of instantaneous formation of rain excess

$$R = p - f \qquad R > 0 \tag{3.23}$$

where R is the rain excess, p the intensity, and f the local infiltration capacity. As was stated, both f and p fluctuate spatially over the watershed (see also chapter .5). The problem is to find the average excess $\overline{R}$. At first sight, it seems that in terms of averages $\overline{p}$ and $\overline{f}$ we have

$$\overline{R} = \overline{p} - \overline{f}. \tag{3.24}$$

Note, however, that wherever $p > f$, $\overline{R}$ builds up with an increase in p; but that where $p < f$ a change in p does not necessarily affect $\overline{R}$. The average as expressed in equation 3.24 overlooks the important, second, part of the definition of equation 3.23. For emphasis we introduce $\overline{R}_e$, the average effective excess rainfall, and retain $\overline{R} = \overline{p} - \overline{f}$. It should be noted that even $\overline{R}_e$ is far too simple a term to describe the real phenomenon. Puddles formed on the soil surface do not necessarily have an outlet. Some of them overflow only after filling up to a certain threshold point. Once overflowing, this threshold may decrease or even disappear by erosion. Also, some of the flowing excess may seep right back into the ground. The simple model is of interest only as a way to demonstrate a possible method required to treat this problem.

In statistical terms

$$\overline{R} = E\left\{R\right\}, \quad \sigma_R^2 = \mathrm{VAR}\left\{(p - f)\right\} = \sigma_p^2 + \sigma_f^2 - 2\rho_{fp}\,\sigma_p\,\sigma_f \tag{3.25}$$

where E is "expected value," VAR is variance, and σ_p^2, σ_f^2 and ρ_{fp} denote the variances and correlation coefficient respectively. One is not interested in $E\left\{R\right\}$, rather, he is interested in the $E\left\{R > 0\right\}$.

Assuming a normal distribution of p and f, R also has a normal distribution with the probability frequency ϕ_R.

$$\phi_R = \frac{1}{\sigma_R\sqrt{2\pi}}\, e^{-\frac{1}{2}\frac{(R - \overline{R})^2}{\sigma_R^2}}. \tag{3.26}$$

A normal distribution is assumed here not because it is likely to occur or be very nearly the right one, but because it serves to illustrate our point. However, in application, the actual distribution must first be identified and the same procedure followed. Similarly, certain independence of fluctuation is inherent in our combined probability calculation which may not always be met by the actual data.

In the special case where $\bar{R} = O$ or $\bar{p} = \bar{f}$, we get from the above distribution function

$$\bar{R}_e = \int_o^\infty \frac{R}{\sigma_R \sqrt{2\pi}} e^{-\frac{1}{2}(R/\sigma_R)^2} dR. \tag{3.27}$$

Changing the variable to $\eta = R/\sigma_R \sqrt{2\pi}$ and integrating

$$\bar{R}_e = \sigma_R \sqrt{2\pi} \int_o^\infty \eta e^{-\pi\eta^2} d\eta = \sigma_R / \sqrt{2\pi}. \tag{3.28}$$

From equation (3.24) for $\rho_{pf} = 0$

$$\bar{R}_e = \left[\sigma_p^2 + \sigma_f^2\right]^{1/2} / \sqrt{2\pi}. \tag{3.29}$$

Hence, obviously $\bar{R}_e \neq 0$ even though $\bar{p} = \bar{f}$.

In general, when

$$\bar{R} = (\bar{p} - \bar{f}) > 0, \tag{3.30}$$

we have (taking into account 3.26)

$$\bar{R}_e = \frac{1}{\sigma_R \sqrt{2\pi}} \int_o^\infty R e^{-(R - \bar{R})^2 / 2\sigma_R^2} dR. \tag{3.31}$$

Changing the variable to $r = (R - \bar{R})$ we have

$$\bar{R}_e = \frac{1}{\sigma_R \sqrt{2\pi}} \left[\int_{-\bar{R}}^\infty r e^{-r^2/2\sigma_R^2} dr + \int_{-\bar{R}}^\infty \bar{R} e^{-r^2/2\sigma_R^2} dr\right]. \tag{3.32}$$

Next, letting $\eta = r/\sqrt{2}\sigma_R$, $\sqrt{2}\sigma_R \, d\eta = dr$ we get

$$\bar{R}_e = \frac{\sqrt{2}\sigma_{\bar{R}}}{\sqrt{\pi}} \left[\int_{-\bar{R}/\sqrt{2}\sigma_R}^{o} \eta e^{-\eta^2} d\eta + \int_o^\infty \eta e^{-\eta^2} d\eta\right]$$

$$+ \frac{\bar{R}}{\sqrt{\pi}} \left[\int_{-\bar{R}/\sqrt{2}\sigma_R}^{o} e^{-\eta^2} d\eta + \int_o^\infty e^{-\eta^2} d\eta\right], \tag{3.33}$$

which may be written

$$\bar{R}_e = \frac{\sqrt{2}\sigma_R}{\sqrt{\pi}} \left[\int_0^{\infty} \eta e^{-\eta^2} d\eta - \int_0^{-\bar{R}/\sqrt{2}\sigma_R} \eta e^{-\eta^2} d\eta\right]$$

$$+ \frac{\bar{R}}{\sqrt{\pi}} \left[\int_0^{\infty} e^{-\eta^2} d\eta - \int_0^{-R/\sqrt{2}\sigma_R} e^{-\eta^2} d\eta\right]. \qquad (3.34)$$

Hence

$$\bar{R}_e = \frac{\sigma_R}{\sqrt{2\pi}} + \frac{\sigma_R}{\sqrt{2\pi}} e^{-\eta^2} \Bigg|_0^{-\bar{R}/\sqrt{2}\sigma_R} + \frac{\bar{R}}{2} + \frac{\bar{R}}{2} \operatorname{erf} \frac{\bar{R}}{\sqrt{2}\sigma_R}, \qquad (3.35)$$

which yields

$$\bar{R}_e - \bar{R}(1 - \frac{1}{2} \operatorname{erfc} \frac{\bar{R}}{\sqrt{2}\sigma_R}) + \frac{\sigma_R}{\sqrt{2\pi}} e^{-\bar{R}^2/2\sigma_R^2}. \qquad (3.36)$$

Here σ_R is as defined in equation 3.25. Obviously R_e is quite different from R. The difference may be expressed as

$$\bar{R}_e - \bar{R} = \frac{\sigma_R}{\sqrt{2\pi}} \exp(-\bar{R}^2/2\sigma_R{}^2) - \frac{\bar{R}}{2} \operatorname{erfc}(\frac{\bar{R}}{\sqrt{2}\sigma_R}). \qquad (3.37)$$

Differentiating $\bar{R}_e$ with respect to R we have

$$\frac{\partial \bar{R}_e}{\partial \bar{R}} = \frac{1}{2}(1 + \operatorname{erf} \frac{-\bar{R}}{\sigma_R\sqrt{2}}). \qquad (3.38)$$

To check equation 3.38, we let $\sigma_R \to 0$, then erf $\to 1$, and $\partial\bar{R}_e/\partial\bar{R} \to 1$, hence $R_e \to R$ as expected when there are no fluctuations in P and f. The same result would be approached for large values of $\bar{R}$; for example, large rains or very small infiltration. The effective addition in rain excess then tends to be equal to the additional rain.

Interestingly, we started with a perfectly linear model. However, because of the simple constraint that $R > 0$, we ended up with a nonlinear case where $\partial\bar{R}_e/\partial p \neq$ constant. According to our model, the line relating R_e to R would be at a slope 1/2 for $\bar{R} = 0$ and would approach 45° for $\sigma_p \to 0$ or for $R \to \infty$ or both. This again indicates that the basic rain-runoff relation is going to have a concave shape that would respond to higher rain fluctuations by having larger runoff.

It is interesting to compare above with the experimental formula suggested by the U.S. Soil Conservation Service (our equation 3.18) where

$$\frac{\partial Q}{\partial P} = \frac{2(P - 0.2d)(P + 0.8d) - (P - 0.2d)^2}{(P + 0.8d)^2}. \qquad (3.39)$$

For small P, $\partial Q/\partial P \rightarrow 0.28/0.64 \sim 0.5$. For large P, $\partial Q/\partial P \rightarrow 1$. The development of this section serves as partial verification of the experimental formula of S.C.S. Other experimental relations have similar forms (25).

Another interesting principle may be demonstrated by equation 3.38. In several places in this report, the use of a differential approach to engineering problems is emphasized. Absolute values of certain functions may be hard to obtain. Moreover, the engineering problem is very often how much a function will change as a result of a change in parameters. In trying to optimize a system, we are also interested in small, incremental changes around the optimum. In equation 3.38 the estimate of $R/\sigma_R\sqrt{2}$ may be very crude. Still, the estimate of $\partial\bar{R}_e/\partial\bar{R}$ will be accurate enough to facilitate the test of small changes in the system. The following is a table of the derivative as a function of $A = \bar{R}/\sigma_R\sqrt{2}$:

A	0.1	0.2	0.3	0.4	0.5	1.0	1.5	2.0
$\partial\bar{R}_e/\partial R$	0.56	0.61	0.67	0.71	0.76	0.92	0.98	0.997

Clearly, when A increases beyond 1.0 there is very little change in the derivative. On a tenfold change in the ratio A from 0.1 to 1.0, the derivative changes only from 0.56 to 0.92.

3.8 Extension of the Statistical Treatment to More Complicated Cases

Having a number of functions Y_i of the variables Q_i, we have the combined frequency of the events $Y_1\ Y_2\ Y_3\ldots Y_n$ as follows:

$$\varphi_y(Y_1, Y_2 \ldots Y_n) = \varphi_Q(Q_1, Q_2 \ldots Q_n)\frac{\partial(Q_1, Q_2 \ldots Q_n)}{\partial(Y_1, Y_2 \ldots Y_n)}. \tag{3.41}$$

To find the average value of a given Y_i, we must weigh it by its frequency function φ_Q and integrate between the physically feasible limits. From physical formulation of processes on a watershed, one should be able to express functional relationships between Q_i and Y_i from which a Jacobian, on the r.h.s. of equation 3.41, would be determined. Once the frequencies φ_Q of Q_i are measurable, the frequency of Y_i can be obtained and average $\bar{Y}_i$ found.

The actual type of Q_i distribution can only be determined by appropriate sampling. Certain types of mathematical expressions of these distributions are more convenient to work with than others. Of special importance is the following property: sums and products of certain individual distribution should produce a new distribution with similar properties. For example, a product of two normal distributions is also a normal distribution. We may thus combine any two variables and instead of specifying them separately by such parameters as their measured averages, variances, and correlations, we can describe them as a combined variable. In that manner, we may choose a combination of variables that has desirable measuring and computational characteristics.

3.9 Some Conclusions and Projections

The purpose in watershed studies is either to have some practical application or to enhance physical understanding. The latter should also be geared, in general terms, to promote more skillful applications.

One extreme approach to watershed studies is to take a certain required output Y_i (equation 3.41) and correlate it with various parameters Q_i. The

most common example is where Y_i is the runoff and Q_i precipitation. Before long it is found that a useful correlation is impossible unless the rain Q_i is broken up into several subparameters or into an actual time and space distribution $Q_i(x, y, t)$. Similarly, one cannot fit any satisfactory curve to include more than one watershed and one storm unless he adds more watershed parameters. Now, these parameters do not necessarily have physical significance. The approach is that of the "black box" (25).

Clearly, such an approach in its extreme has several deficiencies. First, it does not necessarily reveal to the engineer how an intended change in certain watershed features would influence the end result. Second, it gives one the illusion of saving experimental work, while in fact it does the opposite. If the approach is purely that of a numerical correlation, it will produce parameters that by themselves may be meaningless. They can become meaningful only if they are in turn correlated with some other measurable watershed and climatic features. This means, by definition, that almost the same full sequence of measurements has to be repeated in many watersheds. If several subunits in a watershed are arranged in parallel, the separate outputs can be superimposed. However, as often is the case, if the subunits are arranged in series, then there is no way to superimpose the outputs unless physical knowledge is involved.

In that respect, many of the existing correlative models are futile as they do not provide for shortcuts in the investigation of new watersheds. The correlative approach is further limited in its technical advantages by the necessary assumption of linear dependence.

In most present hydrologic studies, an attempt is made to correlate runoff with rain. At best, one can then obtain a prediction power in one sense. An artificial sequence of runoff storms can be produced that will, expectedly, have a certain statistical pattern of peaks, total volumes, and reoccurrences. We may call this form of predicting a power of policymaking. A single event cannot be predicted. Moreover, one cannot expect to understand the mechanisms involved so as to intelligently intervene in nature. If the prediction power is limited to the above, if full research effort has to be repeated for every new watershed (or even for the same one only with somewhat increased or decreased parameters), and if the black-box approach does not reveal physical behavior, then, we had better look for an alternative procedure.

As justification for continuing studies of rain-runoff correlation, it could be argued that many years of rainfall data are available for many locations while runoff data are hard to obtain. However, this contention is questionable because it is not clear how good are the rainfall data and because, as yet, the only places where correlations exist are where runoff measurements are available--in which case we may just concentrate on runoff measurements.

The above is, of course, an oversimplification. There are more reasons to believe that in the future rain could be well measured, and such data could be used for more than just runoff prediction. Moreover, rainstorms can probably be typified for relatively large climatic regions.

The modification in our approach should be in providing more physical insight into the system in place of statistical correlations.

However, a similar danger lies in a radical effort to use strictly physics to the bitter end. There, too, we may end up with an unbelievably large volume of measurements. As we will never be able to rid ourselves from the statistical nature of some of the data, there is a practical limit to the necessary refinement in our physical model.

One can summarize that the approach to watershed problems should involve some approximate physical models, which are not too complicated, and with a feasible way to obtain unknown watershed parameters by measurements. The difficult question is to find the practical optimum. The optimum would change with respect to different practical problems and with the improvements of measuring methods and computational methods.

3.10 A Scheme to Choose an Approach

The analysis (chapters 3 and 4) provide certain tools to determine an approach. First, one should line up the possible practical applications and the available technical implements. They should be specified in several terms: (a) The design quantities D_i; (b) the dependence of these quantities on certain watershed outputs Y_j, as in equation 3.41; (c) the sensitivities $\partial D_i/\partial Y_j$; and (d) the rough nature of Y_j (random or consistent). At the other end, one should line up all possible input parameters $Q_i(x, y, t)$ as in equation 3.41 of the climate topography, soil, and water. A mention should be made of boundary and initial conditions even when they are not exactly known.

The ultimate purpose is to get the Y_js in the range, resolution and accuracy needed to solve for the D_is. One should also decide what type of prediction one wishes to have on Y_j in terms of the length, time projection, space, and form of presentation.

If Y_j is not directly available, easily measurable or otherwise economically obtainable in the proper form, one has to look for some $Y_j(Q_r)$ expression and measure Q_r.

Every relation $Y_j(Q_r)$ can now be obtained in two ways--by the black box approach and by the physical argumentation. A physical formulation brings out new parameters as defined by the Jacobian in equation 3.41 that are useful only if they can be more easily estimated. Every partial derivative in the Jacobian of equation 3.41 defines a certain sensitivity of Y_j to Q_r. Having the tolerance on $\partial D_i/\partial Y_j$, one can determine the tolerance on $\partial Y_j/\partial Q_r$ and on the Q_rs themselves.

Some parameters Q_r and Y_j of low sensitivity factors can be dropped altogether or grouped with some others. Different Q_rs or Y_js can be grouped in some convenient form (submatrices in equation 3.41). Thus, one ends up with a number of possible investigation sequences leading to the required design parameters D_is. If now we attach values (in time, cost, and odds to gain new research and development) to each of the sequences, a strategy may be chosen.

This task, even if partly fulfilled, will probably produce a great savings, will highlight problems that are today neglected, and will prove certain present efforts futile.

Numerous examples can be cited. We shall present only a few. If the raingage measurement can be biased to 10 to 20 percent and if runoff often makes up 10 percent of the measured rain, then 100 to 200 percent errors can occur in predicting runoff. If still, some models are successful in such a prediction some "fudge" factors are hidden in them and they are anything but physically or practically sound.

A model that predicts flow peaks well but does not predict low flows well is totally unfit for estimating water yields but may be perfect for some safety structures (prediction of small storms and large storms, not parts of the same one).

A partial building of a project and a properly planned followup may often produce the most economical way towards a newly designed, more precise solution.

Often, tributary drains have a minimum nominal size and they are, therefore, totally insensitive to hydrological parameters. The hydraulic flow parameters in the drains have a limited accuracy. Drain dimensions are rounded up to the nearest standard size. The required reoccurrences of the design storms are very hard to determine quantitatively. Flows larger than the design storm will be temporarily stored above ground. Combined together those arguments lead to several conclusions: (a) Accurate hydrographs are often not necessary; (b) the engineering solution is insensitive to many watershed parameters; (c) small storage capacity can cheaply increase the safety factor to such an extent that one may save by avoiding intricate investigations; and (d) the part of the runoff hydrograph, which is most needed, is the slope of the peak part (above a certain discharge), to provide for an estimate of the necessary storage.

Existing hydrograph studies have to do mainly with distribution of rain volumes. For the purposes of water quality and pollution, streamline patterns are more important. Erosion studies can hardly utilize the existing channel hydrographs. They call among other things for detailed studies of the hydrological configuration within small subunits of the watershed. (See chapter 7.)

In the following chapter, several ideas will be presented that may be helpful in some approaches. Some are of general applicability, others may prove useless in view of the criteria brought above. As a complete analysis of the type suggested above is as yet not available, our preferences must remain to some extent, intuitive.

Chapter 4 A METHOD FOR AVERAGING WITH PRACTICAL AND PHYSICAL SIGNIFICANCE

4.1 Measurement and Average of a Product

Consider a relation of the form

$$Y = Y(Q_1, Q_2, \ldots Q_n) \tag{4.1}$$

which is obtained by some physical relation at a point in time and space between some output Y and the parameters Q_i. Each parameter Q_i may be by itself a function of some measurable entities. Usually, the r.h.s. of equation 4.1 can be broken into a sum of terms. Each term is a product of several Q_i's. In the following, we shall treat only one such term for example, $(Q_1 \cdot Q_2 \cdot Q_3)$ to illustrate a technique of obtaining meaningful averages.

Consider Y as a measurable entity over an area A and time Δt. For example, it may be the runoff over a plot during a certain time. The measurement is of the average Y, that is, of $\overline{Y}$. Thus, we can transform equation 4.1 into a measurable form by integration over the area A and time Δt. Here, we shall average only over the term (product) $Q_1 \cdot Q_2 \cdot Q_3$. Each other term can be treated in the same way (all terms have the same dimensions). Assuming additivity of Y, we get the arithmetic average:

$$\overline{Y} = \frac{\int Y \, dA \, dt}{\Delta t \; A} = \frac{\int Q_1 \, Q_2 \, Q_3 \, dA \, dt}{\Delta t \; A}. \tag{4.2}$$

To interpret the r.h.s. of equation 4.2 let us assume that we know of a meaningful, measurable average, $\overline{Q}_2$ over area A, depth D and time t, so that averaging Q_2,

$$\overline{Q}_2 = \frac{\int Q_2 \, dA \, dy \, dt}{A \; D \; \Delta t}, \tag{4.3}$$

D being a thickness of the soil profile. We then simply multiply and divide the r.h.s. of equation 4.2 by the numerator on the r.h.s. of equation 4.3. We thus obtain

$$\overline{Y} = \overline{Q}_2 \, \frac{D \int Q_1 \, Q_2 \, Q_3 \, dA \, dt}{\int Q_2 \, dA \, dy \, dt}. \tag{4.4}$$

We may now find it convenient to define $\overline{Q}_1$ as an average of Q_1 weighted by $\overline{Q}_2$:

$$\overline{Q}_1 = \frac{D \int Q_1 \, Q_2 \, dA \, dt}{\int Q_2 \, dA \, dy \, dt}. \tag{4.5}$$

We simply multiply and divide equation 4.4 by the numerator integral of equation 4.5 to get

$$\overline{Y} = \overline{Q}_1 \overline{Q}_2 \frac{\int Q_1 Q_2 Q_3 \, dA \, dt}{\int Q_1 Q_2 \, dA \, dt} = \overline{Q}_1 \overline{Q}_2 \overline{Q}_3 . \tag{4.6}$$

In fact, this defines the type of average $\overline{Q}_3$ required for Q_3. $\overline{Q}_3$ is found as an average, weighted by Q_1 and Q_2 over A and t. In most cases, not all Q_i's can be measured. In the above example, Q_3 is not directly measured. By measuring $\overline{Y}$, $\overline{Q}_1$, and $\overline{Q}_2$, we may then generate the values of $\overline{Q}_3$.

4.2 An Example of Averaging Horizontal Flow Component in a Watershed

Consider the simplified equation for flow through the watershed in an anisotropic profile (the justification for such a formula is given in chapter 6 and is similar to the old "rational" equation for runoff except that it includes a rain dependent coefficient and the land slope).

$$\frac{\partial q_x}{\partial S} \cong p(U - 1) \tan \alpha, \tag{4.7}$$

where q_x is the horizontal flux component, and $\partial q_x / \partial S$ is the added contribution from a section of a flow tube, p the precipitation rate, U the ratio of horizontal to vertical conductivity, and $\tan \alpha$ the surface slope. This equation is used only as an illustration. It is quite possible, for example, that the derivative with respect to S has to be taken on the r.h.s. of equation 4.7 as well (see chapter 6).

We cannot measure the horizontal component of flow q_x at a point. However, we may measure the discharge Q contributed from a stream tube of length L and width B at the main channel. We can integrate q_x over the thickness of the profile $0 \leqslant y \leqslant D$ over drainage length S (total length L), and over stream tube width b and time Δt to get

$$\overline{q}_x = \frac{\int p(U - 1) \tan \alpha \; b(S) \, dS \, dy \, dt}{\Delta t \, D \int b(S) \, dS} . \tag{4.8}$$

Let us assume that p is measurable and we can integrate it over the area $\int b(S) \, dS$ and time Δt. Thus, we get

$$\overline{p} = \frac{\int p \; b(S) \, dS \, dt}{\Delta t \int b(S) \, dS} \tag{4.9}$$

and using the technique suggested in the previous section we divide and multiply equation 4.8 by the numerator of equation 4.9 to obtain

$$\overline{q}_x = \overline{p} \frac{\int p(U - 1) \tan \alpha \; b(S) \, dS \, dy \, dt}{D \int p \; b(S) \, dS \, dt} . \tag{4.10}$$

We can now define the average of tan α. It should be weighted by the local rain intensity p,

$$\overline{\tan \alpha} = \frac{\int p \tan \alpha \; b(S)\, dS\, dt}{\int p\, b(S)\, dS\, dt} \tag{4.11}$$

so that

$$\overline{q}_x = \overline{p}\; \overline{\tan \alpha}\; \frac{\int p(U - 1) \tan \alpha \; b(S)\, dS\, dy\, dt}{D \int p \tan \alpha \; b(S)\, dS\, dt}. \tag{4.12}$$

The average (U - 1) is then automatically defined as

$$\overline{(U - 1)} = \frac{\int p(U - 1) \tan \alpha \; b(S)\, dS\, dy\, dt}{D \int p \tan \alpha \; b(S)\, dS\, dt}, \tag{4.13}$$

so that $\overline{q}_x = \overline{p}\,\overline{(U - 1)}\;\overline{\tan \alpha}$.

There are, of course, other ways to obtain somewhat different operational forms. For example, tan α is a permanent property of the landscape, and we may not wish to weight it by the variable rain, p. Thus, one can exchange the order of equations 4.9 and 4.11 and define first the average slope,

$$\overline{\tan \alpha} = \frac{\int \tan \alpha \; b(S)\, dS}{\int b(S)\, dS} \tag{4.14}$$

and then

$$\overline{p} = \frac{\int p \tan \alpha \; b(S)\, dS\, dt}{\Delta t \int \tan \alpha \; b(S)\, dS}. \tag{4.15}$$

The average anisotropy as in equation 4.13 remains unchanged. The arithmetic average of the rain becomes interestingly irrelevant. For the estimate of runoff, one would weight it by the slope of an areal element b(S) dS.

If the soil properties were known and fixed, one could prefer to further change the order of averaging. First, one would define the average slope, as in equation 4.14. Then the average anisotropy (U - 1) could be defined by

$$\overline{(U - 1)} = \frac{\int \tan \alpha \; (U - 1)\; b(S)\, dS\, dy}{D \int \tan \alpha \; b(S)\, dS}. \tag{4.16}$$

Finally, the average rain is obtained by maintaining that $\overline{q}_x = \overline{p}\,\overline{(U - 1)}\;\overline{\tan \alpha}$ so that

$$\overline{p} = \frac{\int p \tan \alpha \; (U - 1)\; b(S)\, dS\, dy\, dt}{\Delta t \int \tan \alpha \; (U - 1)\; b(S)\, dS\, dy}. \tag{4.17}$$

The product $\overline{\tan\alpha}\ \overline{(U - 1)}\ \bar{p}$ from equations 4.15, 4.16, and 4.17 produces as expected the average horizontal flux, q_x

$$\overline{\tan\alpha}\ \overline{(U - 1)}\ \bar{p} = \bar{q}_x = \frac{\int p \tan\alpha\ (U - 1)\ b(S)\ dS\ dy\ dt}{\Delta t\ D \int b(S)\ dS} \tag{4.18}$$

as originally defined in equation 4.8. In this case the average rain is weighted by the area b(S) dS and by the runoff production capacity (U - 1) tan α. Notably, however, (U - 1) is not a constant and by itself depends on p(t).

The product $\overline{(U - 1)}\ \overline{\tan\alpha}$ may be quite significant from a pedological point of view as is explained in chapter 8. It is quite conceivable that for the purpose of giving soil types hydrological characteristics, one may wish to calculate first $\overline{(U - 1)}\ \overline{\tan\alpha}$ by measuring $\bar{q}_x$ and $\bar{p}$ using simple definitions. It is, for example, possible to define both $\bar{p}$ and $\overline{\tan\alpha}$ simply as in equations 4.9 and 4.14, respectively, so that $\overline{(U - 1)}$ becomes

$$\overline{(U - 1)} = \frac{\int b(S)\ dS}{D \int p\ b(S)\ dS\ dt \int \tan\alpha\ b(S)\ dt} \int p \tan\alpha\ (U - 1)\ b(S)\ dS\ dy\ dt \tag{4.18a}$$

(compare with equation 4.13).

4.3 Average Flux Vector

In the above, q_x was considered to be a scalar as we chose the axis S in the direction of q streamlines. In fact, it is vectorial in nature. By dividing the watershed into subwatersheds, each subunit has an outlet where the flux can be measured and direction assigned to it.

The average flux vector from the watershed as a whole is a vectorial average of the separate flux vectors from each subunit. For a more detailed reasoning see Zaslavsky (68). Measuring the combined flux at the outlet of a watershed consisting of a number of subunits, we may then assign to it a certain overall directional drainage. At the same time tan α is also a vectorial entity as it has an aspect.

4.4 Possible Simplification of the Average by Using Arithmetic Averages and Variances

Analysis in the preceding section indicates that the horizontal flow component q_x may be expressed on the average for a small section as follows:

$$\bar{q}_x = \bar{p}\ \bar{f}\ \bar{\underline{m}} \tag{4.19}$$

where f = (U - 1), m = tan α and p is again the precipitation. Let us say that the averages of the individual parameters are to be written as before in the following order.

$$\bar{q}_x = \frac{\int q_x \, b(S) \, dS \, dt}{\int b(S) \, dS \, dt} \tag{4.20}$$

$$\bar{\underline{m}} = \frac{\int \underline{m} \, b(S) \, dS \, dt}{\int b(S) \, dS \, dt} \tag{4.21}$$

$$\bar{P} = \frac{\int pm \, b(S) \, dS \, dt}{\int m \, b(S) \, dS \, dt} \tag{4.22}$$

$$\bar{f} = \frac{\int pfm \, b(S) \, dS \, dt}{\int pm \, b(S) \, dS \, dt} \tag{4.23}$$

Let us assume also that $\bar{q}_x$ and $\bar{\underline{m}}$ are measurable. We now wish to simplify $\bar{p}$ by using its unweighted arithmetic average $\bar{p}'$ rather than the weighted average of equation 4.22 so that

$$\bar{p}' = \frac{\int p \, b(S) \, dS \, dt}{b(S) \, dS \, dt} \tag{4.24}$$

Setting $p = \bar{p}' + \Delta p$ into equation 4.22, where Δp is the fluctuation around the unweighted average we get

$$\bar{P} = \bar{P}' + \frac{\int \Delta pm \, b(S) \, dS \, dt}{\int m \, b(S) \, dS \, dt}. \tag{4.25}$$

Now, the same substitution can be used for $m = \bar{m} + \Delta m$ where equation 4.25 reduces to

$$\bar{P} = \bar{P}' + \frac{1}{\bar{m}} \frac{\int \Delta p \Delta m \, b(S) \, dS \, dt}{\Delta t \int b(S) \, dS}. \tag{4.26}$$

If there is a correlation between p and m (as will be shown in section 5.8 there is), the correction term on the r.h.s. of equation 4.26 does not vanish. If there is a linear dependence between Δp and Δm, then the fluctuations become proportional to $(\Delta m)^2$ and the correction term of equation 4.26 reduces to the F.R.I. of sections 3.1 and 3.2.

In a similar way it can be shown that equations 4.13 and 4.13a approximate each other. In fact, when the fluctuations are small or the correlations are small all the averages approach the unweighted arithmetic ones.

4.5 The Differential Approach

In the above, it has been indicated how one may obtain meaningful averages. For making the discussion plausible, examples from watershed hydrology have been analyzed. While the specific examples pose some interesting

new possibilities, they are by no means well established. The averaging process was in fact an integration over a subsystem and the definition of a continuum (68). The continuous averages can then serve to produce differential expressions. Many inferences may be obtained by using differential expressions of the continuous parameters. One example was described at the end of section 3.7. Here another example is cited for flow of groundwater into an auger hole.

The rate of flow Q is given by

$$Q = GK(\Phi_1 - \Phi_2) \tag{4.27}$$

where Φ_1 and Φ_2 are any two potentials measured at certain predetermined points, K the hydraulic conductivity, and G some geometrical function. Equation 4.30 is already a macroscopic or "averaged-out" formula. Even if the soil is not uniform and not isotropic, G will represent an average or equivalent coefficient to be used in such a simplified model. (GK) is really defined as $Q/(\Phi_1 - \Phi_2)$. This is the philosophy behind pumping tests and many other measuring techniques. If we treat a case of a limited range of Q so that GK does not change much, equation 4.30 can be improved to cover nonlinear cases by simply expressing the differential of Q and redefining GK as a partial derivative,

$$dQ = GK\, d(\Phi_1 - \Phi_2),\ GK = \frac{\partial Q}{\partial(\Phi_2 - \Phi_1)}. \tag{4.28}$$

This is more than a fancy. It has some fundamental and practical significance. Often one cannot actually measure absolute values of our parameters but we can introduce or measure small changes. Moreover, often in practice, all we ask is to predict the outcome of such changes.

The main conclusion is that there is a way to incorporate physics to any degree of refinement and still maintain the average sense of the entities. There is some freedom in choosing some types of averages but there is no freedom to choose them arbitrarily. The arithmetic average can sometimes be used with a correction for the variance and correlation of the different parameters.

Chapter 5 SOURCES OF HETEROGENEITY AND THEIR IMPLICATIONS

5.1 Introduction

Chapters 3 and 4 showed that the heterogeneity of the hydrologic system must be a part of the formulation even when average parameters are considered. The heterogeneity may appear in several ways. There is a degree of randomness in the spatial and temporal distribution of different inputs such as the rain, wind, and temperature. In that respect an individual event is unpredictable. However, in a statistical sense, one may describe frequencies, expected values of averages, variances, autocorrelations, cross correlations, and confidence limits of some input parameters. For any given purpose, one can define how to obtain a useful accuracy.

A given output may be by itself of a random nature (like drainage flow). It may become predictable in details, only in two senses. First, like the input it may have statistics. Secondly, an output in one event can be predicted if the input for the same event is known. In this last case, one has to know also the detailed description of the system (soil and water) that transforms the input into the output. This transmitting system is also heterogeneous.

The output may be more or less smooth, depending most often on the delay time with respect to the period of the characteristic fluctuation. This marks one of the more fundamental problems in choosing the size of ensemble or time period over which to average the data. Averaging over periodic parameters and over a limited time and space will produce a new statistical population whose statistics depend on the size of periods or space and the range over which the averaging is made. The input is composed by a superposition of periodic phenomena with frequencies ranging from a fraction of a second all the way to centuries or millennia.

There is still another type of statistics besides the inputs and the transmitting system. Even if one assumes a continuous variation of the hydrologic parameters, the sampling or measurement is usually spotty in nature. The hydrologic system poses one of the most difficult systems to treat statistically or otherwise. Thus we would be wise if we at least try to eliminate the statistics of spotty measurements as much as we can, to reduce the effort of measurement, and to render more significant results. Some measurements of this kind are already being used. A pumping test for groundwater investigation is such a measurement.

As long as we have no physical relations between the different parameters, each one presents a separate problem for statistical inferences. Physical relations may hint towards a functional relation between parameters and suggest the proper form of measurement. At the same time such relations lend themselves to an intelligent intervention in the system even when not all the parameters are measurable.

In the following, we shall mention briefly some sources of heterogeneity and some of their possible hydrologic significance.

5.2 Heterogeneity and Nonuniformity

An attempt will be made to define the terms "heterogeneous and nonuniform." To do that one must first consider the fluctuation and distribution of a certain property in time and in space. To understand the significance

of these definitions, one has to consider the changing parameters of the system as having various frequencies and amplitudes.

As an example, we may consider density of a porous material that varies spatially. On the smallest scale, we have atoms and molecules that affect the density with fluctuation between zero and extremely high values. On a larger scale elementary soil particles, assuming all are of the same density, pose a frequency modulation with a constant amplitude equal to half the density of a particle. Aggregates pose still a larger scale with a smaller frequency and a smaller amplitude, if we average over the elementary particles. Smaller frequencies or larger cycles can be observed as we average over small parts of a field. The field may then be found to be nonuniform since its average properties vary from one place to another. The rain is certainly nonuniform with time. Every year we get a different rain. Taking a long range average, the rain in a certain region is said to be heterogeneous between years. The average monthly rain is nonuniformly distributed throughout the year, but it is heterogeneously distributed within a month, which is our averaging period. In a given storm we may consider the average hourly precipitation. The hourly average rain is said to be nonuniform throughout the storm. But it is heterogeneous within the hour. Five minute rain gages will indicate nonuniform rain distribution within the hour. However, if shorter measurements could be taken, we would probably find high-rain intensities of a cycle length of a part of 1 minute. In fact, a single raindrop is the smallest source of storm heterogeneity both in time and space, posing the highest frequency and highest amplitude.

We may conclude, therefore, that we always measure and manipulate some average properties with a finite range of averaging in time and space. Focusing our attention on a certain time and space range, fluctuations of a higher frequency (ranges smaller than the range of our averaging) are considered as heterogeneities. The change of our averaged properties with time or space must be considered as nonuniformities, relevant to the problems at hand.

Obviously these two terms are relative. The distinction between heterogeneity and nonuniformity is made by the choice of scale or by the method of measurement, or the range of our samples or measuring implements. The heterogeneity is described by statistical modes. The nonuniformity and its exact distribution in time and space is a part of our functional formulation of the problems. Measuring on a given scale, we may suspect the existence of heterogeneity but know only about nonuniformity. To account for the heterogeneity, we have to add observations at a smaller scale.

If we wish to study the details of a soil profile (our measuring implement is only a few centimeters in size), we would say that the profile is nonuniform. We still have to remember that each horizon in the soil is also heterogeneous as it is made of particles and aggregates. The heterogeneity will influence and control the apparent physical laws such as the flow of water. However, in describing the flow, we use average, continuous description that can be only nonuniform.

The scale of observation can always be decreased to include more details in our profile description. This amounts to the transfer of heterogeneity into nonuniformity. The limitation is of a practical nature. The choice of scale poses one of the most difficult theoretical and methodological problems in the study of soil and water.

5.3 Definition of a Continuum

A continuum is useful in applying available physical concepts and mathematical techniques. Following Zaslavsky (68), the continuum can be formed by averaging over subsystems. One subsystem has a volume V_1 and a second subsystem V_2 is slightly translated with respect to the first one so that it almost overlaps (fig. 5.1) the first. We now attach average values to the centers of the subsystems. These centers form a continuum. By increasing the size of the subsystem the average properities at the centers become smoother.

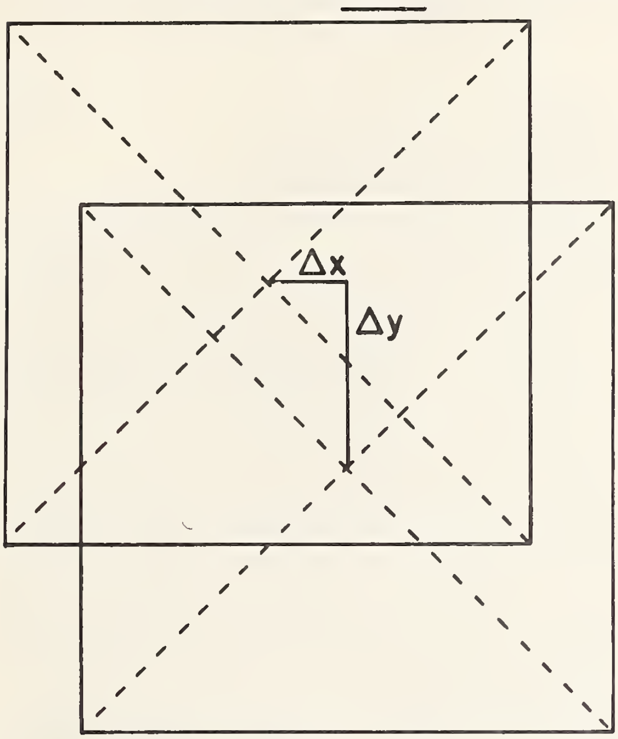

Figure 5.1.--Subsystems for definition of continuum.

In a discontinuous system, differentiation is not possible but integration is. Averaging amounts to such integration. It has been practiced by soil physicists and hydrologists for many years. The discrete nature of some physical phenomena also led mathematicians to develop the use of generalized functions and functionals that are in effect averages over the function space with some weighting functions (24).

When the sources of heterogeneity are known the size of the subsystem can be chosen so that the errors involved in neglecting fluctuations of certain frequencies will be kept down to a useful limit. The upper size of the subsystem should be such that it will not smooth out changes that are of interest. In any case, infinitesimal translations of the subsystem will produce infinitesimal changes in the averages.

5.4 Interflow or Quick Return Flow

A system analysis comparing the rain input P(t) and the runoff output Q(t) seemingly indicates several components that contribute to the runoff (25). There is a very fast response with high peaks and a short delay time. This is usually attributed to surface or overland flow that eventually reaches stream channels. A typical delay time is 30 minutes; for example, on watershed 94W in Coshocton, Ohio.

A tail flow in the channel can be attributed to shallow ground water drainage. Its delay time is several hours from a watershed of a few square miles. This component is usually attributed to temporary rise in perched water horizons and shallow soil drainage into the stream. It is called an interflow or a quick return flow. A third component has a very large delay time of weeks or months and small amplitudes. It is attributed mainly to ground water flow, also called base flow. There is a degree of arbitrariness in dividing the output into these three superimposed components. The intention

of this section is not to invalidate the existence of the interflow as found from the outflow analysis, nor to approve of it. The only purpose is to indicate the consequences of adopting the concept of such an interflow component.

A simple calculation would show that in most cases a shallow drainage to the mainstream would contribute a negligible quantity of water. As a rough estimate consider a gradient i, an hydraulic conductivity K, and flow depth D. The flow per unit length of stream Q is then roughly as follows:

$$Q = DKi. \tag{5.1}$$

If the total length of the channel is s and the flow is symmetrical

$$V = 2\ DKis, \tag{5.2}$$

where V is the rate of interflow through the channel boundaries.

Let us take a very large $i = 0.1$ and a high $K = 10^{-3}$ cm./sec. Assume D to be 50 cm.; thus, to contribute an interflow of 1ℓ. per second the length of the drain s must be on the order of 1 km. If the contributing area is 500 x 1000m.^{2} then this rate would be the equivalent to about 0.2 mm./day rain. If the "interflow" is caused by a rise of 50 cm. in water table or by perched water of this depth, then a very long drain is necessary for a relatively small runoff discharge. In other words, to explain the rates that are commonly associated with "interflow," the drain density must be very large. This implies the following conclusion: one may wish to maintain the concept of "interflow" and use it in its physical sense as an explanation for a significant part of the runoff. In such a case, he must also conclude that it occurs not only in the main channels but in a much denser system of dents, rills and gullies. As their density increases the possible gradient i towards them becomes larger. This is the only way larger runoff intensities of fast response, in relatively flat soils, may be accounted for.

This very simple physical consideration implies the importance of the microrelief. We shall later bring out more reasons to amplify the importance of soil surface heterogeneity. However, already, the question is raised here whether it is possible to make a distinction between water that flows above the soil or in the soil. As with several other entities, it becomes a matter of scale. Focussing our attention on a certain channel size, all smaller water conduits become a part of shallow water flow that includes water that is below the surface, at the surface, and above the surface. Conduits of a smaller scale become part of the heterogeneous nature of the landscape (section 5.2).

5.5 The Search for an Overland Flow

The usual picture drawn for the runoff mechanism is roughly as follows. The rate of rain exceeds the infiltration capacity of the soil. Then an incipient storage of water is formed on the soil surface. Subsequently, a laminar flow starts towards the nearest runoff channel. This overland flow increases as a function of the slope and as a function of temporary storage above the soil surface. Such a flow regime is yet to be observed. Rather, the common picture is that soil seems to be unflooded in certain areas and flooded in others, in the form of puddles or small rivulets. Evidently, the mechanism of water

concentration into free channel flow starts at the bottom of slopes, at V-shaped indentations of high frequency and concave dents several centimeters to a few meters apart. If this is the case, then our physical approach and our watershed description must also start on this scale. The physical and statistical nature of microrelief should appear in our hydrologic macroscopic formulation. We cannot hope to have an actual description of all the details of the microrelief in every watershed. Yet the microrelief of the field is probably the one single factor most responsible for the relation of input to output and should be given some overall values for parts of the landscape. An attempt to intercept surface flow would be futile unless it is aimed at a definite, well-defined channel. We may conclude that surface flow can be unambiguous only if it is synonymous with a flow in a channel of large enough size.

5.6 The Influence of Slope on Water Concentration

Assume that the volume of horizontal outflow Q due to a slope i is proportional to it and to a characteristic time Δt.

$$Q \sim i\Delta t. \tag{5.3}$$

At this time, the amount of rail P is proportional to the length of the slope L_i and the time Δt thus

$$P \sim L_i\Delta t. \tag{5.4}$$

The degree of concentration of precipitation on the slope towards the stream is

$$\frac{Q}{P} \sim \frac{i}{L_i}. \tag{5.5}$$

The slope itself is usually steeper for shorter lengths of slope. In fact, very short slopes may have the steepest slope. Thus, one may roughly say

$$\frac{Q}{P} \sim \frac{\epsilon_i}{L_i^2} \tag{5.6}$$

where ϵ_i is the amplitude of soil surface irregularity, $1/L_i = f_i$ is the frequency of a certain slope i, and

$$\frac{Q}{P} \sim \epsilon_i f_i^2. \tag{5.7}$$

This derivation is very qualitative and only a rough indication of a possible trend. But, as before, it points out the relative importance of the high frequency, small scale irregularities in producing localized concentration of water.

The mechanism of water concentration during the storm is obvious when the rain intensity exceeds the infiltration capacity and a staturated zone is formed below and above the soil surface. However, as will be shown (chapter 6), the concentration occurs in the unsaturated state as well.

5.7 Observation of Total Runoff Versus Total Rain

In a smooth horizontal plane or a very slightly sloping one, runoff will start only when the rain exceeds the infiltration capacity. If the rain and infiltration are uniform over an area, the runoff must start everywhere at the same time. Rains of a small rate, of a short duration, or of a small total volume, would not produce any runoff. (Kraijenhoff 25) indicated that some runoff occurs even for small rains. This can be explained only if the rain is not uniform, the infiltration is not uniform, or there are some fast changes in the surface water distribution (see especially section 3.7). There must be a combination of these three factors that will produce a significant rain excess at certain field localities. Some water, although concentrated, will remain as temporary surface storage that will seep in at a later time (locally, or by flowing over unsaturated sections of the soil surface). The probability of interconnections between such localized puddles will affect the output discharge from the watershed.

In short, we have to look for some mechanisms by which runoff will start much before the average rain intensity exceeds the average infiltration capacity. In line with our concepts of heterogeneity (section 5.2), we shall specify below some more aspects of the heterogeneous rain and microrelief. In addition, we shall consider the heterogeneity of the soil profile as an important parameter.

There must be another clarification of the concept runoff. Usually, runoff is considered as the amount of water that appears in a channel outside the soil. The artificial distinction between surface runoff and interflow implies that within the watershed some water is flowing inside the soil. However, this water is not measured unless it appears outside the soil. Evidently, it is possible that runoff will exist even when water never appears above the ground. To have a physically consistent definition of runoff, we should consider the runoff as the flow component which is either horizontal or parallel to the soil layers. This definition does not exclude runoff from, or through unsaturated soil.

5.8 Additional Parameters in Rain Measurement and Soil Microrelief

Assuming that a rain gage is measuring the flow through a horizontal plane, we have to make a correction for the actual rain distribution on the soil surface. Consider, a rain that is falling at an angle β with the vertical (fig. 5.2). A rain gage records P_o. The horizontal projection of the soil slope s is ($s \cos \alpha$) and the amount of rain on a horizontal soil surface would be ($P_o s \cos \alpha$), where α is the slope of the soil surface. The horizontal equivalent of the sloping soil $s \sin \alpha \tan \beta$ and the additional amount of rain would be $P_o s \sin \alpha \tan \beta$. Thus, if the rain gage measures P_o, the sloping surface will receive P units of rain, where

$$P = P_o (1 \pm \tan \alpha \tan \beta). \tag{5.8}$$

The sign is obviously positive for α as shown in figure 5.2 and negative should the slope be in the opposite direction

The same rain intensity will have different effects depending on α and β. The local rain concentration can be more than double the recorded one. Without accounting for this effect, to predict runoff in its traditional sense

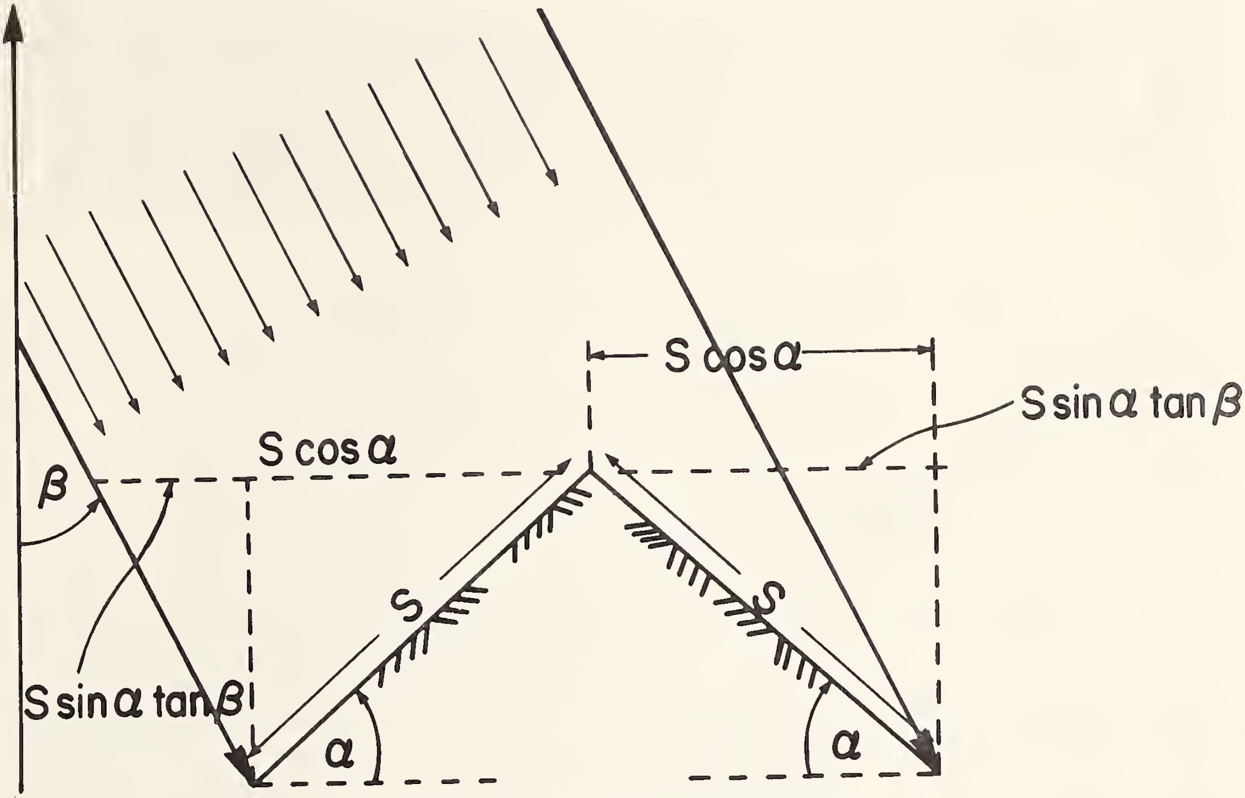

Figure 5.2.--Slanting rainfall onto microrelief.

as a rain excess is not possible. Usually, the overall slope will not exceed a few percent but local microrelief fluctuations of soil surface can have very large slopes.

A general analysis requires considering the orientation of the soil slope with respect to the rain (fig. 5.3). Taking the prevalent slope as one direction and the horizontal projection of the rain path as another, we define an angle γ as in figure 5.3. Equation 5.8 then changes to

$$P = P_o(1 \pm \tan\alpha \tan\beta \cos\gamma). \tag{5.9}$$

As an example, let us consider a cultivated field with furrows of slope $\tan\alpha = 0.7$, where the rain slope is 45^o and normal ($\gamma = 0$) to the furrow lines ($\tan\beta = 1$, $\cos\gamma = 1$). Thus

$$P = P_o(1 \pm 0.7),$$

when the 45^o rain is parallel to the furrow lines ($\gamma = 90$).

$$P = P_o\ .$$

Obviously, changes in local rain intensity due to rain direction may be as much as 70 percent. The variance of the rain would be around 0.5 (chapter 3). It is clear that with the measurements of only average rain and infiltration rates, one cannot predict runoff.

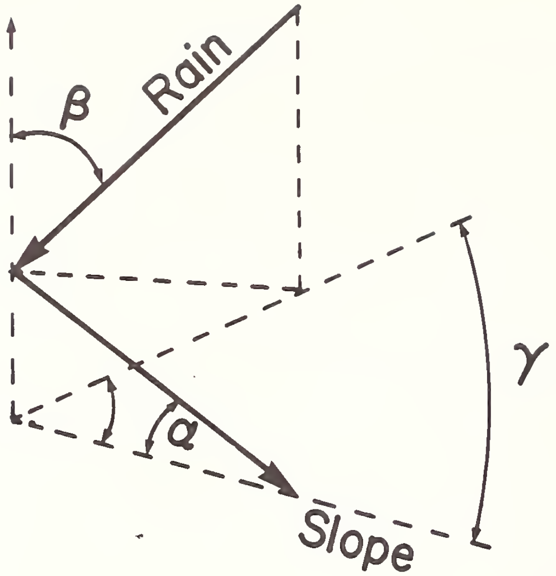

Figure 5.3.--Definition of angle γ for rain at angle β falling onto slope of angle α.

To analyze an area according to equation 5.9 let us pick an arbitrary direction γ_0 (that is $\gamma_0 = 0$ northward). We define ζ as an angle that the slope of a certain areal element makes with γ_0 and ϵ as an angle that the rain projection on the horizontal plane makes with γ_0. Then

$$\cos \gamma = \cos (\zeta - \epsilon) = \cos \zeta \cos \epsilon + \sin \zeta \sin \epsilon \tag{5.10}$$

and we get

$$P_\epsilon = P_0 \left[1 + \tan \alpha \tan \beta (\cos \zeta \cos \epsilon + \sin \zeta \sin \epsilon)\right]. \tag{5.11}$$

A certain relief is specified by $\tan \alpha$ values and its direction by $\cos \zeta$ and $\sin \zeta$. Average rain concentration can be found by specifying the values of ϵ and β. This must be introduced as a soil parameter and as a rain parameter to any formula that would undertake prediction of runoff as a rain excess by physical reasoning. Even if rain measurements are not satisfactorily accurate, the above microrelief parameters and rain aspects can and should be utilized. Slanting rain is a very common experience; also some slopes get wet and shed water before others.

5.9 Heterogeneity and Erosion

At least one important mechanism of erosion is by raindrop splashing. Positive protection of soil surface often proves effective in reducing erosion drastically changing the runoff significantly. Thus, at least in such cases, erosion by splashing is implied. One cannot conceive of erosion by splashing without frequent enough water conduits that will capture the splashed material and carry it away. We may conclude that the microrelief is greatly responsible for concentrating the water as a vehicle for eroded material and in providing locally converging slopes that would tend to concentrate splashed material.

The interaction between erosion and hydrology deserves a deeper investigation. Rills and gullies that make the microrelief are themselves a product of hydrologic processes. It seems that splashing above would smooth them out as it can throw material mainly downward. Thus, other mechanisms must be looked for. As rills and gullies often occur at points of very small water velocities, there must be some mechanisms other than erosion by the drag forces of a water stream in stream channels. This problem will be discussed in chapter 7.

5.10 Heterogeneity of Soil Profile

The soil profile is often layered. This layering may be caused by history of deposition, by metamorphosis under stresses, or by secondary pedological development of profile differentiation.

Whether one should try to describe these layers in detail as profile nonuniformities is a practical question. The alternative is to consider the profile as a whole. Then, the layering becomes an anisotropy. In treating a watershed as a whole and in trying to produce flow equations for it, the details must be considered as heterogeneities that are averaged. In measuring only inflow and outflow from a certain area, we determine the scale of our measuring implement as much coarser than these details.

5.11 A Highly Permeable Layer at the Very Surface of the Soil

The definition of a porous medium (4) is related to a multiphase system of some finite size so that average properties can be studied. To use the mathematics for a continuum, one must consider subsystems of finite size (68) and relate their average properties to their center of gravity section 5.3. The continuum is obtained by translating the subsystem slightly and thus obtaining translated centers of gravity and smoothly changing variables (fig. 5.1). Alternatively, in producing the continuum, one has to use the weighted integrals of certain properties over a volume (see chapters 3 and 4). This definition is more than just a mathematical fancy. We are not really capable of considering a porous medium in any other way. The problem of a surface storage water and its distinction from soil water is really no exception. Where is exactly the boundary between soil water and overland flow? It depends to a great extent on particle size, roughness of the soil surface, and the scale of observation.

More specifically, could anybody decide whether a pore or a depression at the soil surface is a part of the air above the soil or really a pore within the soil? Rather, one can observe a transition. If the soil is very smooth and the particles are very small, there will be a relatively sharp transition from the soil porosity n_s to the air porosity n_a which is 100 percent. The horizontal conveyance of water through the transition zone will be small. The conductivity to water above the soil will depend on the thickness of water film. In fact, the discharge is proportional to the third power of the film thickness. If any irregularity in soil surface is of thickness ϵ, then the transition between soil and air will be on the order of ϵ and the water transmissivity through this transition zone and parallel to it is proportional to $\epsilon^3 i$, where i is the slope. This is before any free water will really be seen without the help of a magnifying glass. When the soil surface becomes rough or soil aggregates are large, this water conveyance at the surface can become larger.

The saturated hydraulic conductivity of the transition from soil to air is changing from that of soil to a very high one by several orders of magnitude. In this transition zone, the stream lines must bend away from the vertical towards the direction of the slope (see sections 6.3 and 6.5). In the transition zone, even when unsaturated, the hydraulic conductivity is never less than the rate of rain p, as gradients are of the order of one. Thus, one can assume that there is a horizontal sheet flow which is at least of the order of ip, where i is the overall slope of the area. However, because of local variation, there will be local flows with a much larger slope, sometimes reaching gradients of a few tenths and up to unity. These will certainly cause local water concentration that may well surpass the local infiltration capacity. The actual embodiment of this flow need not be through saturated flow, but may be by unsaturated flow and by splashing of raindrops.

In conclusion, a mechanism is demonstrated that leads toward localized surface concentration of rain. A far more important conclusion is that the arbitrary division between soil water and surface water is meaningless. Rather, it seems that there is more or less a gradual transition between the two. One can say that the saturated hydraulic conductivity of the soil profile is always variable and is extremely high at the very soil surface.

5.12 An Example of Some Heavy Soils in Israel

An extensive drainage research was conducted on grumusols (like volk clays in Texas) in Israel. When well cultivated, the soil had the following features. The drainable porosity changed with depth. It has less than 1 percent at a depth of 1.5 ft. and nearly 20 percent at what could roughly be called soil surface. The total porosity changed in similarly way from some 50 percent at the 1.5-foot depth to nearly 70 percent at the soil surface. The weighted hydraulic conductivity of the top 1.5 ft. was about 10^{-1}cm./sec., while at a depth of about 3.0 ft. it was 10^{-6} cm./sec. or less (fig. 5.4).

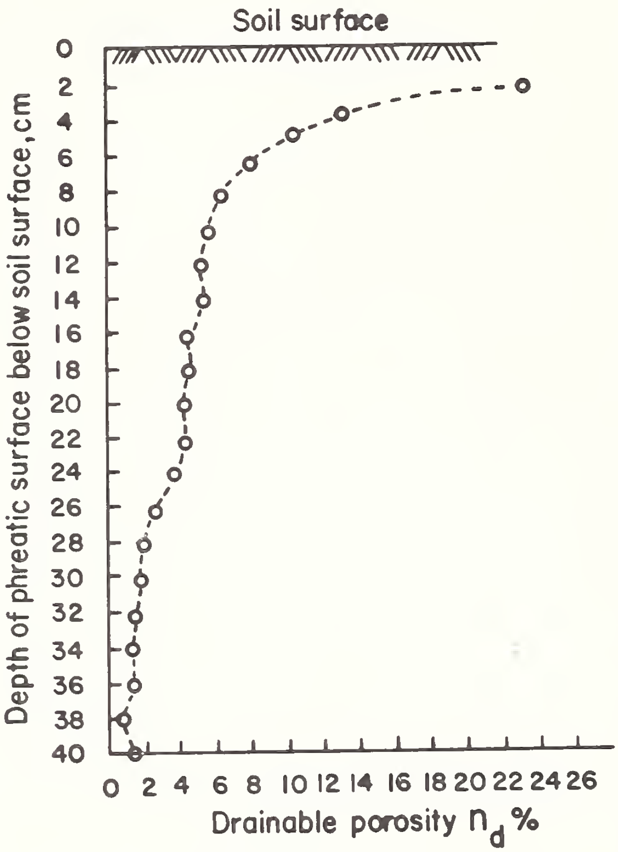

Figure 5.4.--Porosity with depth on a grumusol.

After 10 days of continuous irrigation of 4 in. of water per day, no obvious runoff occurred as the drainage system accounted for all excess water. A shallow ditch of some 0.5 feet was flowing full and intercepted practically all the water that should have been drained through a nearby drainage tile. Practically all the water went through some shallow flow. It was impossible to decide when soil flow stopped and runoff started.

Of course, this would not be the case in every soil. However, something of the kind must exist. The effects on streamflow will probably depend on the ratio of rain intensity to conductivity of the impeding layer, on the frequencies of soil surface irregularities, and on their amplitudes.

Although the picture is already complicated, some concept of a surface anisotropy should be interjected here. A soil that is plowed in a certain way may demonstrate anisotropic hydraulic conductivity for flow parallel to the soil surface. Anyone who has observed furrows overrun by water has probably seen how at times the water opens a route which is not parallel to the furrows and not parallel to the slope vector. In other words, the surface conductivity has a different value in two orthogonal directions.

In considering the soil profile as a whole, it is possible to have a conductivity tensor that has three different component values in three directions.

5.13 An Estimate of the Shallow Conductivity and Storage

The size of our subsystem should be several times larger than the amplitude of a soil surface irregularity at least in one or two directions. How large depends on how detailed we wish the study to be, that is, what irregularities we wish to draw into our average features and what irregularities we wish to specify in our detailed boundary conditions.

Studying a single furrow, the size of our subsystem should probably be several times that of a single aggregate; for example, the aggregate size 10 mm. and the thickness of the subsystem at least 10 to 20 mm. This means that the transition of porosity and hydraulic conductivity from that of the soil bulk to that of the air will be through at least 20 to 40 mm. (fig. 5.5). Surface

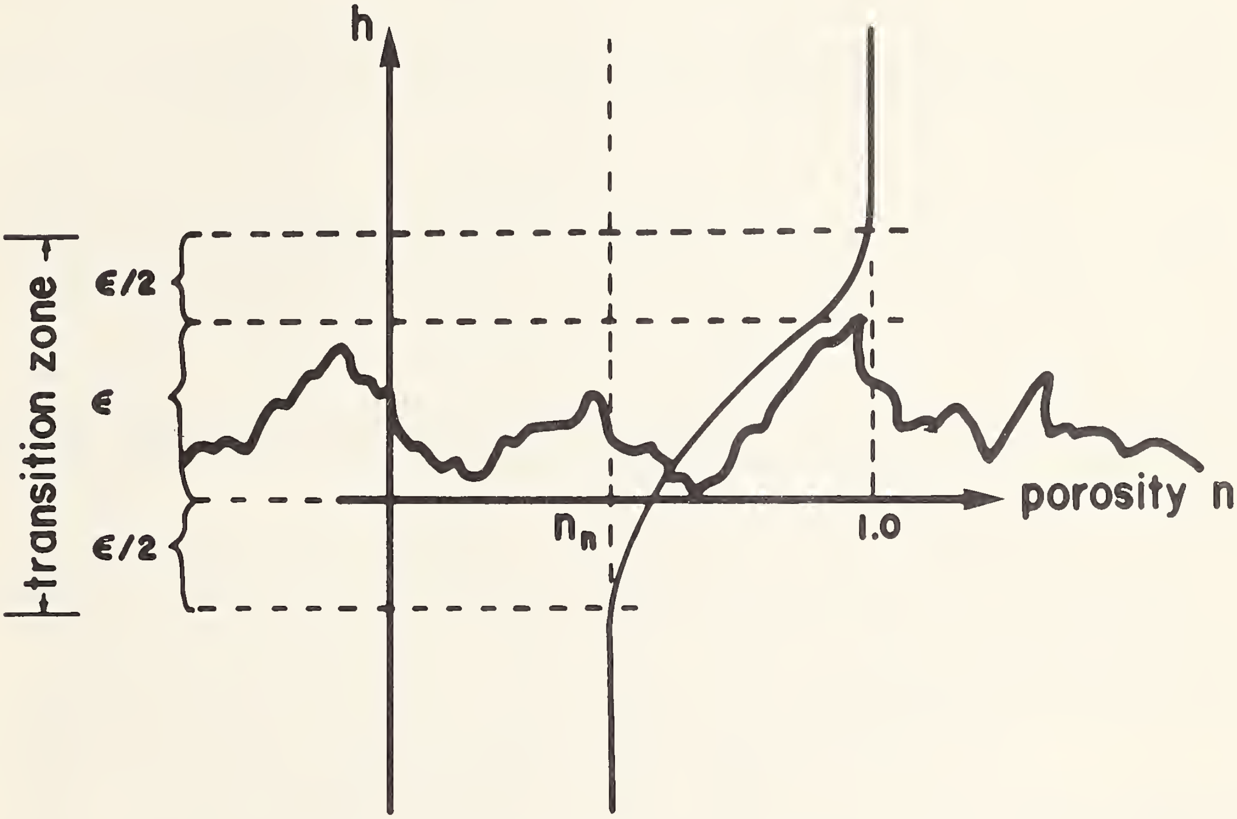

Figure 5.5.--Transition of porosity n across soil surface.

flow in its fullest sense will not occur until water surface is more than 5 to 10 mm. above the average position of soil surface or above the top of the aggregates. This means that we have a shallow flow of 10 to 20 mm. thickness with porosities ranging for example from 50 to 100 percent and hydraulic conductivity ranging from that of the soil bulk to some very high values. In fact, the lateral transmissivity T changes with height h as follows, h being measured from some depth within the soil:

$$dT = Kdh, \tag{5.12}$$

where K is the hydraulic conductivity. If we assume a laminar flow with an effective depth D(h) (yet unknown $D < h$), with a discharge Q of the form $Q = \alpha i D^3$ (i being the slope and α some proportionality coefficient that may vary with h) then

$$dQ = 3\alpha i D^2\, dh \tag{5.13}$$

$$dT = \frac{dQ}{i} = 3\alpha D^2\, dh. \tag{5.14}$$

We may wish to use the formula (4, equation 9.10):

$$K = K_r \left[\frac{n - n_o}{n_r - n_o}\right]^3 \tag{5.15}$$

where K_r is some reference hydraulic conductivity at a reference porosity N_r and N_o a stagnant porosity (N_o may be near field capacity). If n varies linearly as

$$n = [h(1 - n_r) + \epsilon n_r]/\epsilon, \tag{5.16}$$

where h is measured from a depth of 0.5 ϵ below the soil surface and ϵ is the size of an aggregate, the minimum added storage through the transition zone is

$$\int_0^{\epsilon} n\, dh - n_r \epsilon \cong \epsilon \frac{1 - n_r}{2}. \tag{5.17}$$

Taking, for example, $\epsilon = 10$ mm. and $n_r = 0.5$, the added storage (before conventionally defined surface runoff takes place) is 2.5 mm. of water. We can similarly calculate K by using equations 5.15 and 5.16:

$$K = K_r \left[\frac{(h/\epsilon)(1 - n_r) + n_r - n_o}{n_r - n_o}\right]^3 = K_r \left[\frac{(h/\epsilon)(1 - n_r)}{n_r - n_o} + 1\right]. \tag{5.18}$$

Where $h = \epsilon$ we find for $n_r = 0.5$ and $n_o = 0.4$

$$K = K_r \left[\frac{0.5}{0.1} + 1\right]^3 = K_r(6)^3 = 216K_r. \tag{5.19}$$

It may be more proper to change K by changing n_o also; that is, decreasing it. Let us assume that n_o decreases linearly from a certain value to zero at the top of the transition zone. At the surface we shall have

$$K = \left[\frac{(1 - n_r) + n_r}{n_r - n_o}\right]^3 K_r \tag{5.20}$$

or with the above values

$$K = K_r \left[\frac{0.5}{0.1} + \frac{0.5}{0.1}\right]^3 = 10^3\, K_r. \tag{5.21}$$

The truth is probably somewhere between these two values $216K_r$ and $1000K_r$. Using equation 5.13, we find at the top of the transition zone

$$3\alpha D^2 = 216K_r, \tag{5.22}$$

$$D = \left(\frac{216K_r}{3\alpha}\right)^{1/2}. \tag{5.23}$$

This is the equivalent depth of film flow when runoff starts. The added transmissivity is of the order of $65K_r$ in cm.^{2}/sec. for the above example. In other words, the transmissivity added at the top 1 cm. transition is equivalent to that of a 65 cm. of soil bulk.

These derivations are no more than illustrations indicating possible ways to treat the problem of the transition zone and estimating its contribution. Other assumptions and other forms of K - n dependence may be more useful.

In the air above the ground, there is a higher concentration of raindrops at any given instant. This is caused by splashing. After a fashion, one may consider this as a highly permeable layer above a less permeable one. Surely, the overall result of such splashing is what would appear to be a sheet flow just above the ground. It seems that this has to be incorporated in the same concept of shallow flow in an anisotropic soil.

A very important feature of this top layer shallow flow is its nonsteady effect. The speed of flow there is much higher than in the soil bulk. Thus, it may affect the output hydrograph significantly. Typical velocities may be hundreds of times the usual flow velocities in the bulk. The shorter the slope stretches, the larger will be the proportion of water that will drain in a short time. Again, one realizes that the high frequency irregularities in soil surface (here really intermediate in size) may be responsible for a large part of the runoff.

When local irregularities are absent and ϵ is not excessive (moderate smooth slopes) the runoff will be smaller or delayed. Similarly, large, high frequency storms should lead to high runoff. Where we have closed furrows that would not allow flow, we have a very large storage (ϵ is very large).

If ϵ is 200 mm. and n_r is 0.5, then we have an extra storage of 150 mm. In this artificially produced medium, n_o is very large when the furrows are closed every few feet. The soil surface can be considered anisotropic with extremely small hydraulic conductivity normal to the furrows.

An interesting aspect of the above is that the rate of shallow runoff will adjust itself to the intensity of the storm. The higher the intensity and length of the storm, the larger will be the participation of the top permeable layer in conveying downhill flow. This is parallel to common features of unsaturated flow where often the rates are adjusted to the input with very little changes in the hydraulic gradients.

Chapter 6 FLOW IN A LAYERED OR ANISOTROPIC SOIL

6.1 Introduction

Chapter 5 showed that every soil has a more permeable layer, however thin or thick, at its surface. To determine when flow in the soil bulk ends and runoff starts is impossible. Runoff is redefined as the flow of water in a horizontal direction or parallel to the soil surface. From a practical point of view, we wish to dispose of too many details in our watershed description. Thus, one is tempted to treat the soil as anisotropic rather than as layered (4, pp. 68-75). The quality of the rain measurements and the random nature of rain make their use for runoff estimates questionable in many cases. Estimating is practically impossible when the runoff is considered as an excess of rain (over and above infiltration and storage). For any reasonable degree of accuracy, the mean rain as is obtained today has far too large a variance. Moreover, there is no assurance against consistent bias resulting from the measuring method, placement of stations, averaged fluctuations, and wind. However, high correlations between rains are found at neighboring locations.[1] They diminish with distance and time. This is encouraging, as the rain measurements could be used as an index for a limited area. Although their absolute values may be incorrect, they still would serve as a yardstick. This is only on condition that their absolute accuracy is not needed for subtraction or addition with other random parameters. In summary, we shall try to produce an approach that considers the landscape as a complete entity, anisotropic and nonuniform. As will be shown in this chapter, the possible use of the rain as a proportional yardstick for runoff becomes a byproduct of this model.

The ability to explain new phenomena is not necessarily a proof for the validity of a theory. However, the crop of corollaries produced by the new model is encouraging. In chapters 7 and 8, we shall relate the theory to some erosion processes and soil forming processes.

The text will lean heavily on the knowledge of flow through porous media, especially in the unsaturated state. Although credit should be given to many scientists, I shall generally refer to a text by Bear, Zaslavsky, and Irmay (4) for its convenience. They present more proper and complete documentation.

6.2 Some Useful Observations in Unsaturated Flow

In the following, I shall draw very briefly some simple rules or observations concerning unsaturated flow. They will be recorded here without proof for the convenience of future reference.

Consider a uniform soil with a saturated hydraulic conductivity K_0. After placing the saturated soil in a column and permitting drainage, the soil in the column will eventually reach an equilibrium condition. At that stage, the pressure at the bottom outlet would be zero ($p = 0$). At any elevation above it, the water pressure head $\pi = p/\gamma$ would be negative (γ representing the unit weight of water).

[1] Hershfield, D. Personal communication.

By supplying a certain water rate I at the top (volume per unit area per unit time), eventually there would be a steady state flow uniformly throughout the column. π will be uniform through most of the column if $I < K_0$, so that π is negative and the soil is not saturated. Only at the bottom there will be a transition from $\pi < 0$ to $\pi = 0$ at the very outlet.

By increasing the rate of flow $I \rightarrow K_0$, the column will become more saturated and $\pi \rightarrow 0$. At $I = K_0$ the soil will be waterlogged and incipient water storage will appear at its top.

It is instructive that $\partial\pi/\partial z = 0$ (z the elevation) except at the bottom and thus the gradient of the head ($\Phi = z + \pi$) becomes a unity. Thus, at all times, the unsaturated $K < K_0$ adjusts itself to the rate so that $K = I$. (See fig. 6.1).

If a more permeable layer is placed above a less permeable one the problem becomes more complicated. One thing can be stated very generally. Assuming that, at the suction prevailing at the interface between the two layers, the bottom layer is still less permeable, the pressure is going to increase towards the less permeable layer (fig. 6.2).

If the bottom, less permeable layer drips into the air (or to a very coarse gravel layer), $\pi = 0$ at the very bottom outlet. The pressure buildup will be more like curve 4. At a lower rate the pressure buildup will look like curve 4'.

If, however, there is a third layer at the bottom that is more permeable (fig. 6.2) one gets a series of curves 1 to 6 with an increasing rate I, as described below.

Curve 1	$I < K_2$:	classically no runoff, I < infiltration capacity.
Curve 2	$I = K_2$:	the beginning of perched water.
Curve 3	$I > K_2$:	but $< K$ and K_3: perched water and saturation parts of layers 1 and 2.
Curve 4	$I < K_1$:	but $I > K_2$ and $I = K_3$: perched water 2 and 3; more exactly layers 2 and 3 are pressurized but with the possibility of entrapped air bubbles.
Curve 5	$I_5 =$	$(D_1 + D_2 + D_3)/\left(\frac{D_1}{K_1} + \frac{D_2}{K_2} + \frac{D_3}{K_3}\right)$: incipient water accumulation appears at soil surface.
Curve 6	$I > I_5$:	runoff in the classical sense is formed.

For a short rain that never reaches steady state, incipient ponding will occur only if $I > K$. For sufficiently long rain to produce ponding, it must be large compared only with K_2 and the thickness of the top layer must not be large.

That the flow regime depends not only on the soil matter itself but on the profile as a whole. Change of boundaries or of water addition procedures may change it completely.

Any rate of infiltration into the soil if continued long enough will reach a steady state (zero flux is a special case).

A less permeable layer at the surface will produce an unsaturated flow underneath (for example curve 3, fig. 6.2). It is sufficient to disturb slightly the soil surface by raindrop pounding or by flooding to produce this effect.

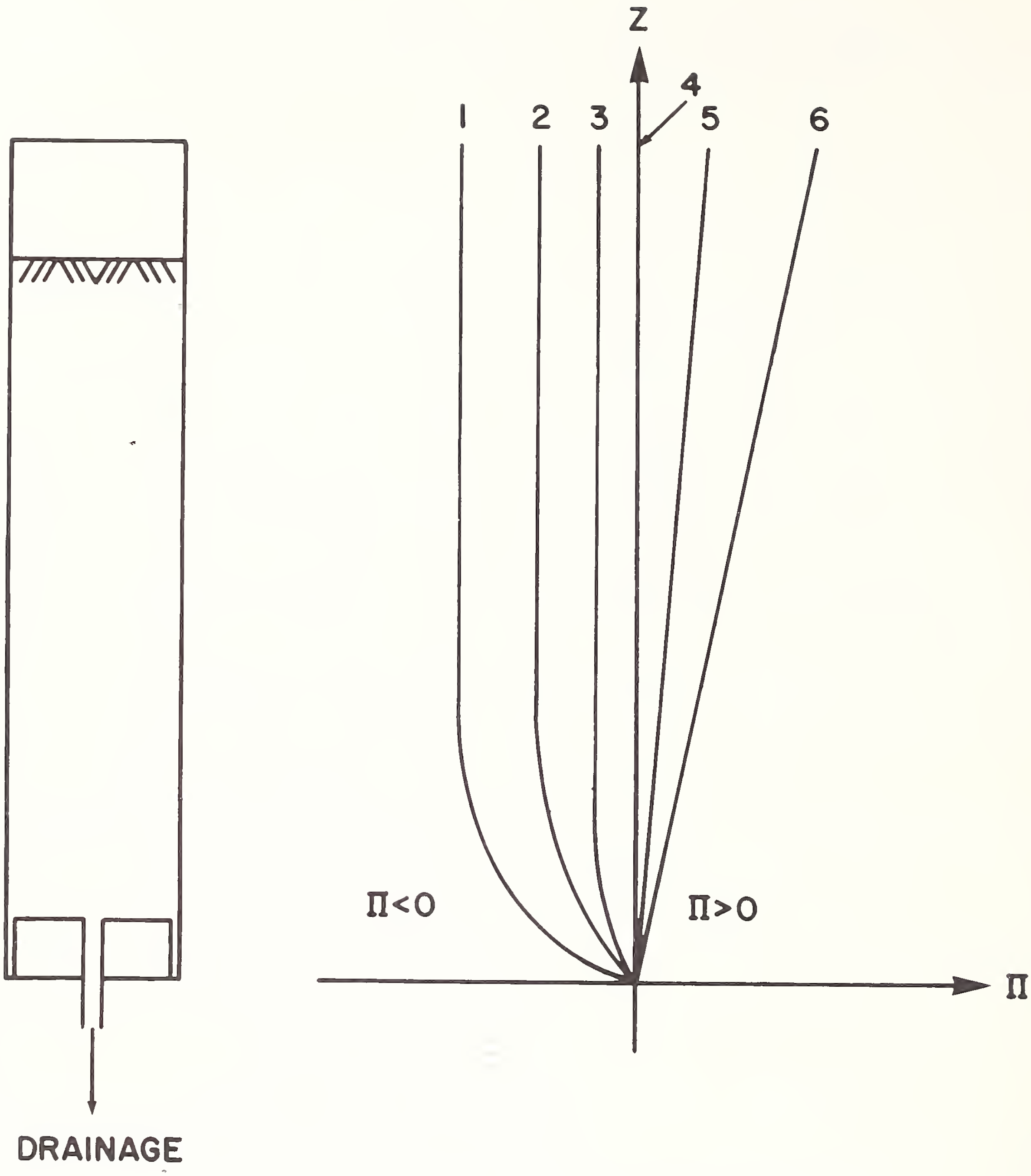

Figure 6.1.--Soil water suction as a function of depth for a draining column.

This does not exclude the existence of a more permeable layer at the very surface above the crust.

Near the water source, conditions will become steady soon after the start of the flow. This is despite the nonsteady conditions farther away.

Many complications are avoided here for the sake of simplicity. Once our concepts are established any number of refinements is possible.

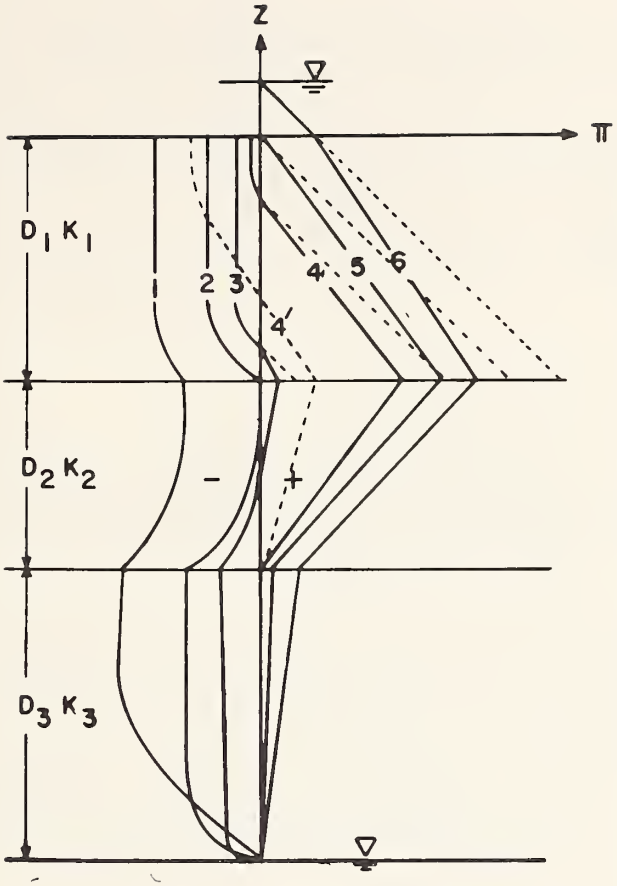

Figure 6.2.--Effect of hydraulic conductivity on pressure distribution in three-layered system.

6.3 Demonstrating "Runoff" in an Unsaturated Soil

The following is not a part of our suggested hydrologic formulation. It is intended only for some microphysical insight as to the why and how.

Consider a more permeable layer above a less permeable one. The soil is sloping and so are the layers. Imagine vertical partitions inserted into the soil that allow only vertical flow. The pressure distribution in two neighboring sections is shown in figure 6.3. Now, removing the partitions there will be formed a horizontal flow component. Both sections are under suction $\pi < 0$. However, comparing neighboring points of the same level (line AB) $\pi_1 > \pi_2$ at the transition zone. As π increases uphill in a horizontal direction flow will tend to be downhill. This lateral gradient occurs where the hydraulic conductivity is the highest in the profile. Gradients in the opposite direction occur as well, but they are in zones of much smaller hydraulic conductivity.

In a uniform soil there will be no lateral flow. In a horizontal soil also there will be no lateral flow. In a sloping layered soil, there will be runoff soon after the rain starts and even before perched water or complete flooding is formed. It has been shown in sections 5.11 to 5.13 that every soil, with no exception, has a more permeable layer at its surface (even above a surface crust). Thus, every soil may produce runoff from any rain. The runoff does not always reach a main channel. The transition zone, in the more permeable layer where π increases, becomes thicker as the rate of rain increases and, the thickness of the zone where there exist horizontal gradients increases. Moreover, with increased π, K also increases. Hence, one may qualitatively conclude that the lateral flow will increase as the rate of rain p or infiltration I increases (as in the order of curves 1 to 6 in fig. 6.2). It seems that the above is a possible model by which the rate of runoff will be a function of time, the rate of rain, and the slope. The runoff is a transient process that starts at a very low moisture content and gradually increases with the rain all the way to a flooded soil. There is no way to determine where the runoff in its old sense starts. Moreover, there is no need for such a distinction.

It may be noted here that the above analysis assumes a steady state. At initial stages of the rain, the soil is relatively dry and there exists a wetting front where the pressure change would lead to an uphill flow. The flow direction will then tend to move towards a line normal to the soil surface.

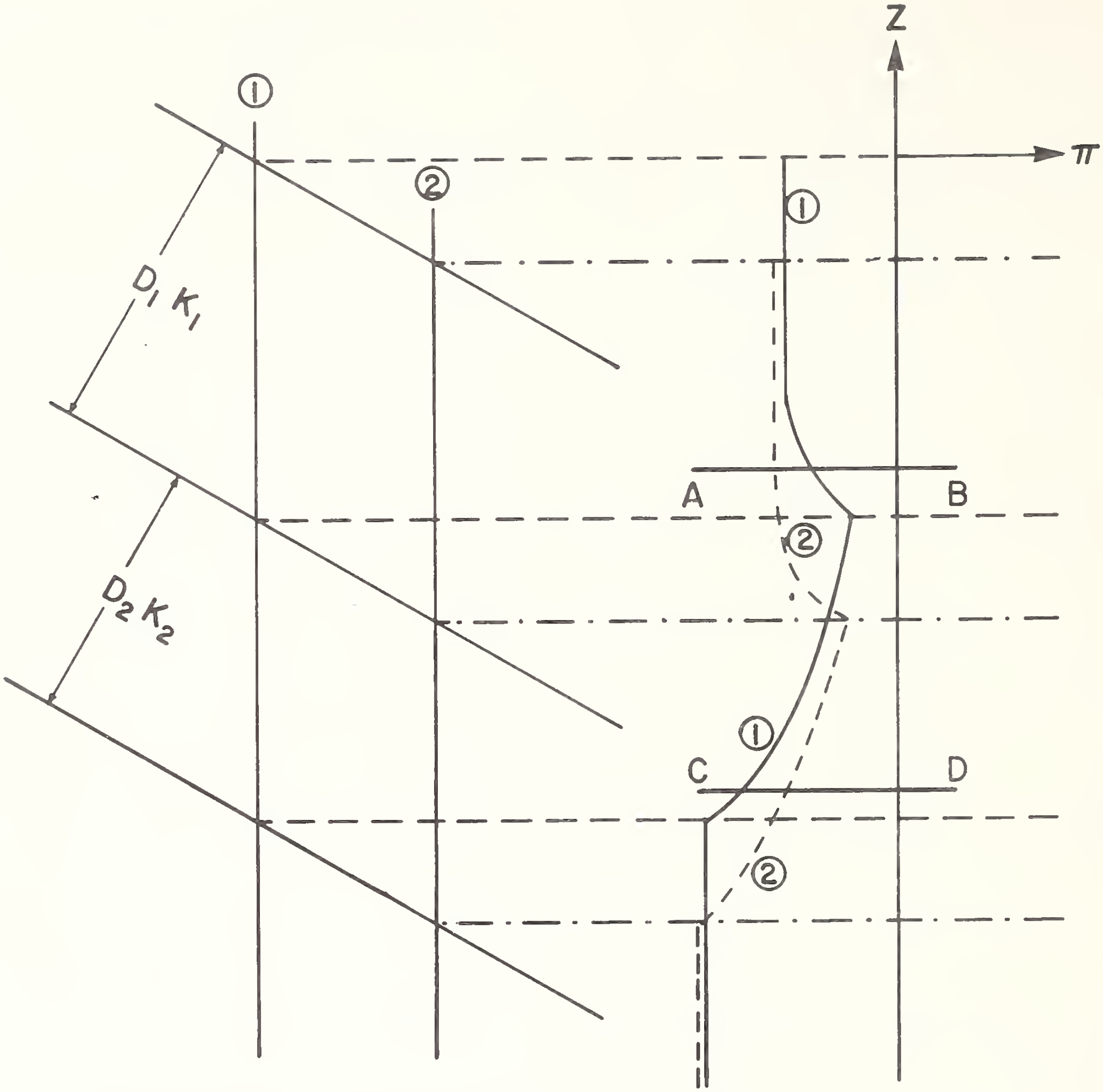

Figure 6.3.--Pressure distribution for flow through a two-layered, sloping soil system.

However, if a crust is formed, or a less permeable layer near the surface, downhill runoff will start soon after the rain because the conditions above the crust will quickly tend to become steady. Small surface irregularities will then produce high slopes and a large degree of water concentration. The above qualitative analysis may be applied to other problems of unsaturated flow, such as evaporation or water redistribution after the rain has ceased.

6.4 Attempts to measure Runoff and Interflow

Because of lateral flow, water will tend to accumulate at concave parts of the slope and at points of transition to a heavier soil. At such points, the vertical, downward rate of flow will increase and so will the lateral one. There will come a point where the soil may actually be flooded or water will concentrate in a stream. This flooding caused by a little local additional

rain or even just by seepage out of the soil. It must however be emphasized that everywhere above these points of accumulation the runoff exists but not necessarily in a saturated state. If $\pi < o$ and the soil is unsaturated, liquid water cannot seep out of the soil. In this case, an attempt to measure the runoff in a drain or an open trench is futile. Moreover, the very attempt to measure will change the boundary conditions so completely that some saturated flow will seep back in and disappear and some unsaturated flow will build up and produce outflow in a totally unexpected place. Many experiments in concrete flumes, with or without a leaky bottom, may be equally meaningless (14).

6.5 Curving of Streamlines Downhill

When a streamline comes from a less permeable to a more permeable layer, the streamline deviates from the normal (fig. 6.4). In uniform saturated layers, this leads to the well-known relationship

$$K_1/\gamma_2 = \tan\alpha_2/\tan\alpha_1,$$

where α_1 and α_2 are the angles between the streamline and the normal in the respective layers of conductivity K_1 and K_2.

This relation may be shown to hold also in the more general case of gradually varying conductivity, in both saturated and unsaturated flow. If the hydraulic conductivity increases gradually, the streamlines would curve gradually away from the axis formed by grad K. Assuming that the rain enters vertically at the soil surface of a slope, it may be proved that

$$\tan\gamma = \frac{K}{K_\alpha}\tan\alpha_1 \qquad (6.1)$$

where γ is the angle between the streamline and the normal at any depth. K_α is the hydraulic conductivity (unsaturated at the very surface) and K is any other hydraulic conductivity within the soil. For a deep enough uniform permeable layer, $K_\alpha = I$. At the point where $\pi = 0$, or more exactly at saturation $K = K_s$. We thus have at a perched water table within a uniform layer

$$\tan\gamma = \frac{K_s}{I}\tan\alpha. \qquad (6.2)$$

Interestingly at that point the ratio between the horizontal flow component q_x and the vertical one q_z is $\tan(\gamma - \alpha)$, so that

$$\frac{q_x}{q_z} = \frac{\left(\frac{K_s}{I} - 1\right)\tan\alpha}{1 + \frac{K_s}{I}\tan^2\alpha}. \qquad (6.3)$$

Within the zone of $\pi > 0$ the flow direction remains the same if K_s is uniform. If the rain p penetrates the soil vertically so that initially $q_x = 0$, then the vertical component q_z must remain the same while q_x varies. Thus we may assume near the surface, where flow is almost steady $q_z = I = p$.

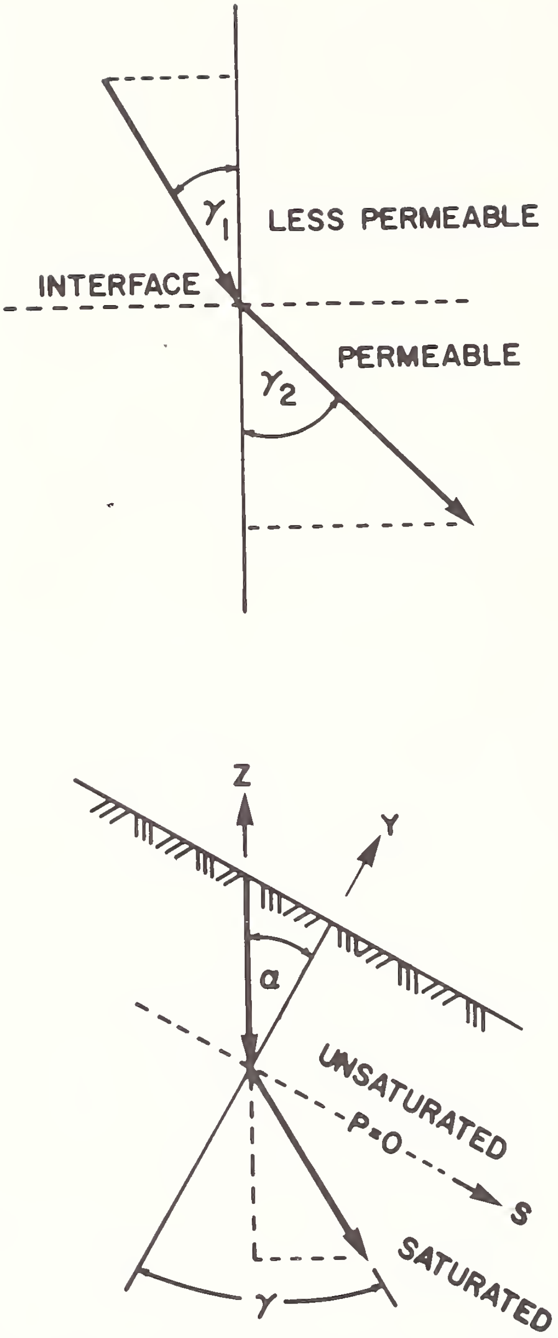

Figure 6.4.--Bending of streamlines at interface between (a) cones of differing conductivity and (b) saturated and unsaturated zones.

We may choose another, rotated, set of coordinates with q_s parallel to the soil surface and q_y normal to the soil surface. It may be shown that in these coordinates $q_s = q_y (K_s/I) \tan\alpha$. The definition of runoff can be either q_s or q_x.

A previous section shows that the hydraulic conductivity at the very soil surface can be several hundred times larger than in the bulk. Thus, one can expect at the very surface $K_s \gg I$ even for extremely high rain intensity. Thus, as soon as the phreatic surface approaches the surface transition zone, the horizontal component becomes very high.

The above analysis indicates that the flow will be slanted downhill. The horizontal component will increase in proportion to the vertical flow component and in proportion to the slope. The extent of the transition zone will be related to K_s/I and to K_1/K_2.

Using equation 6.1 one can actually compute the shape of the streamlines by relating K to a certain π and solving for π (y) by using methods such as those of Bear, Zaslavsky, and Irmay (4, sections 9.15-9.18).

As a possible illustration, consider a long uniform slope with a uniform horizontal component. Because of water conservation, there must be

$$q_z = I \tag{6.4}$$

so that at $\pi > 0$ within the more permeable layer

$$q_x = I \frac{\tan\alpha \left(\frac{K_s}{I} - 1\right)}{1 + \frac{K_s}{I} \tan^2\alpha} \tag{6.5}$$

Taking the partial derivative $(-\partial q_x/\partial x)$ of the r.h.s. of equation 6.5, one can learn how changes in $\tan\alpha$, K_s and I can cause buildup of moisture in various parts along the slope. If the slope reduces with x (a concave slope), q_x/I decreases. This can be either by a decrease in q_x or by an increase in I or both. An increase in I is associated with higher ponding and increased q_x. Thus, we should expect that at higher slopes and at more curved

transitions from a higher slope to a lower one, or in a horizontal transition from higher to lower K_s there would be ponding, overland flow, and sometimes seepage out of the soil.

6.6 The Soil as an Anisotropic Medium

In the following, we shall treat anisotropy in a very limited sense. We simply exchange the layered soil of some unknown thickness D by an anisotropic one. The flow equation is then

$$q_i = \sum_j K_{ij} F_j , \tag{6.6}$$

where q_i are the flux components and F_j are components of the hydraulic gradient. K_{ij} are the conductivity component where $K_{11} \neq K_{22}$. We shall make reference to two coordinate systems. One is (x, z), where x is the horizontal and z the vertical. The other is (S, y), where S is parallel to the layers and y is normal to the layers. The angle between S and x or between y and z is α, the slope of the land. We may write our equations in either one of these systems. The (S,y) system is really the set of principle axes where the terms K_{ij} for $i \neq j$, vanish. Thus

$$q_s = - K'_{11} \left(\frac{\partial z}{\partial S} + \frac{\partial \pi}{\partial S}\right) = - K'_{11} \left(- \sin \alpha + \frac{\partial \pi}{\partial S}\right)$$

$$q_y = K'_{22} \left(\frac{\partial z}{\partial y} + \frac{\partial \pi}{\partial S}\right) = - K'_{22} \left(\cos \alpha + \frac{\partial \pi}{\partial y}\right), \tag{6.7}$$

where K'_{11} and K'_{22} are the conductivities parallel and normal to the soil layers, respectively, and $\tan \alpha$ is the slope of the land.

We may now express the ratio $-q_y/q_s$ which is the angle between the flux vector and the soil surface $\tan \beta$ from equation 6.7 (fig. 6.5)

$$\tan \beta = - q_y/q_s = \frac{1 + \dfrac{1}{\cos \alpha} \dfrac{\partial \pi}{\partial y}}{U \left[\tan \alpha - \dfrac{1}{\cos \alpha} \dfrac{\partial \pi}{\partial S}\right]} \tag{6.8}$$

$$U = K'_{11}/K'_{22} = \text{anisotropy}$$

In many cases $\partial \pi/\partial S$ is negligible. However, for small enough $\tan \alpha$, a positive $\partial \pi/\partial S$ may be produced at some horizontal soil transition so that β may become negative--seepage out of the soil. Neglecting $\partial \pi/\partial S$ and $\partial \pi/\partial y$ we get

$$-q_y/q_s = \tan \beta \sim (U \tan \alpha)^{-1}. \tag{6.9}$$

Equation 6.9 and its more general form equation 6.8 are very significant. The direction of flow will become more nearly parallel to the soil surface as the slope, $\tan \alpha$, increase and the anisotropy U increases.

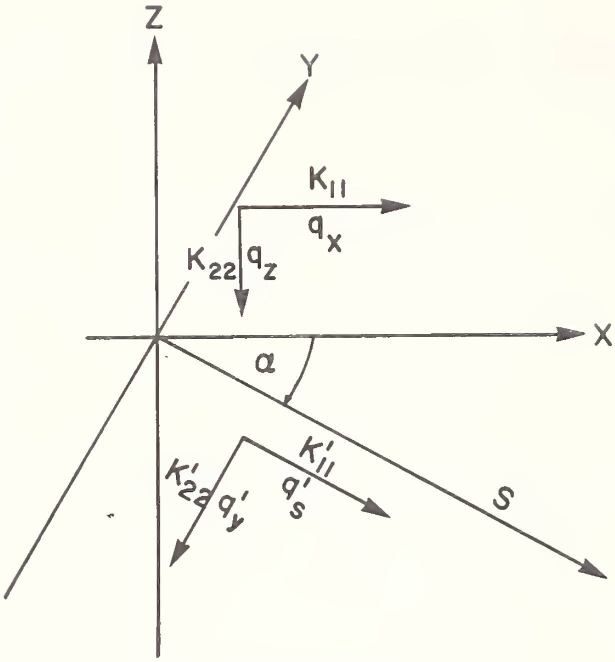

Figure 6.5.--Translation of axes for anisotropic soil.

As shown above, the parallel conductivity increases as q_y increases or as the rain extends in time and intensity. Thus, U is a function of the flow rate itself. At some very small rates of vertical flow, $U \to 1$. At high rates of vertical flow caused by high intensity rain or water accumulation, U would increase. U will reach its maximum when the field is flooded and a large part of the flow takes place at the soil surface. We may write

$$q_s \cong q_y \, U\,(q_y) \tan\alpha \qquad (6.10)$$

at a given point in the profile q_s is one possible expression of the runoff. It is a continuous function of the rain.

In equation 6.8 $\partial\pi/\partial y$ diminishes after prolonged infiltration. $\partial\pi/\partial S$ is never very large unless there is a sharp horizontal change in conductivity. A simple test shows that when the flow is practically uniform $\partial\pi/\partial S = 0$ and when the field is waterlogged so that $\partial\pi/\partial S = \sin\alpha$, q_s is zero, that is, there is no horizontal flow. In every other case there is runoff q_s.

It may be more proper to express the infiltration q_z and runoff q_x in the vertical and horizontal coordinates. By a simple transformation

$$\begin{aligned} q_x &= q_y \sin\alpha + q_s \cos\alpha \\ q_z &= q_y \cos\alpha - q_s \sin\alpha \end{aligned} \qquad (6.11)$$

and from equation 6.7

$$\begin{aligned} q_x = - K'_{22} \cos^2\alpha \;\Bigg[&\left(1 + \frac{\partial\pi}{\partial y}\,\frac{1}{\cos\alpha}\right) \tan\alpha \\ &- U\left(\tan\alpha - \frac{\partial\pi}{\partial S}\,\frac{1}{\cos\alpha}\right)\Bigg] \end{aligned} \qquad (6.12a)$$

$$\begin{aligned} q_z = - K'_{22} \cos\alpha \;\Bigg[&\left(1 + \frac{\partial\pi}{\partial y}\,\frac{1}{\cos\alpha}\right) \\ &+ U\left(\tan\alpha - \frac{\partial\pi}{\partial S}\,\frac{\tan\alpha}{\cos\alpha}\right)\Bigg] \end{aligned} \qquad (6.12b)$$

$$-\frac{q_z}{q_x} = \frac{1 + \dfrac{\partial\pi}{\partial y}\,\dfrac{1}{\cos\alpha} + U\left(\tan^2\alpha - \dfrac{\partial\pi}{\partial S}\,\dfrac{1}{\cos\alpha}\right)}{\left[\dfrac{\partial\pi}{\partial y}\,\dfrac{1}{\cos\alpha}\right] \tan\alpha + U\,\dfrac{\partial\pi}{\partial S}\,\dfrac{1}{\cos\alpha}} \,. \qquad (6.13)$$

Under proper conditions the above can be approximated as

$$\frac{q_z}{q_x} = \frac{1 + U \tan^2\alpha}{(U - 1) \tan \alpha} \tag{6.14}$$

This ratio describes the angle between the horizon and the flux. Note the similarity between equations 6.14 and 6.3. Now, the real runoff q_x can be expressed at any given point as a function of the vertical infiltration q_z:

$$q_x = - q_z \frac{(U - 1) \tan \alpha}{1 + U \tan^2\alpha} \tag{6.15}$$

In a steady state flow $q_z = - I$, if $I < K$, K being the minimum in the profile. Thus, we may approximately write

$$q_x = I \frac{(U - 1) \tan \alpha}{1 + U \tan^2\alpha} \tag{6.16}$$

As before, one must remember that U itself is a function of I. An actual solution is at least as difficult to get as to solve an unsaturated flow problem with time variable boundary conditions. However, many qualitative conclusions and useful models may be obtained. For example, for small values of $\tan \alpha$ and moderate U values, one may find

$$q_x \sim I (U - 1) \tan \alpha. \tag{6.17}$$

Assuming a uniform I

$$\frac{\partial q_x}{\partial S} = I \left[\tan \alpha \frac{\partial}{\partial S} (U - 1) + (U - 1) \frac{\partial}{\partial S} \tan \alpha \right]. \tag{6.18}$$

Clearly, if $\frac{\partial}{\partial S}(\tan \alpha)$ is negative--that is, a concave landscape--one of two things must be true. It is possible that $\partial q/\partial S$ is negative, namely, there must be some water accumulation and outflow from the soil. Another possibility is that $\partial S (U - 1)$ is positive. This means that if on a flatter surface there is no considerable water outcrop, then the degree of anisotropy must increase. An increase in the anisotropy $(U - 1)$ is also related to waterlogging. As will be shown in chapters 7 and 8 this phenomenon may be very significant in soil forming processes and in erosion.

6.7 Some General Formulas and Derivations

In the principal axes S and y with the flow potential

$$\Phi = z + p/\gamma = z + \pi \tag{6.19}$$

$$q_y = - K'_{22} \; \partial\Phi \; \partial y \tag{6.20}$$

$$q_s = - K'_{11} \; \partial\Phi/\partial S, \tag{6.21}$$

where S is along a streamline in an isotropic soil surface. In the system x and z (horizontal and vertical)

$$q_z = - K_{21} \; \partial\Phi/\partial x - K_{22} \; \partial\Phi/\partial z \tag{6.22}$$

$$q_x = - K_{11} \; \partial\Phi/\partial z - K_{12} \; \partial\Phi/\partial z. \tag{6.23}$$

To transform from the primed (principal) system K'_{11} and K'_{22} to any other, one can use the Mohr's circle fig. (6.6)

$$K_{11} = K'_{11} \cos^2 \alpha + K'_{22} \sin^2\alpha =$$

$$= \frac{1}{2} (K'_{11} + K'_{22}) + \frac{1}{2} (K'_{11} - K'_{22}) \cos \alpha \tag{6.24}$$

$$K_{22} = K'_{11} \sin^2\alpha + K'_{22} \cos^2\alpha =$$

$$= \frac{1}{2} (K'_{11} + K'_{22}) - \frac{1}{2} (K'_{11} - K'_{22}) \cos \alpha \tag{6.25}$$

$$K_{12} = K_{21} = \frac{1}{2} (K'_{11} - K'_{22}) \sin^2 \alpha. \tag{6.26}$$

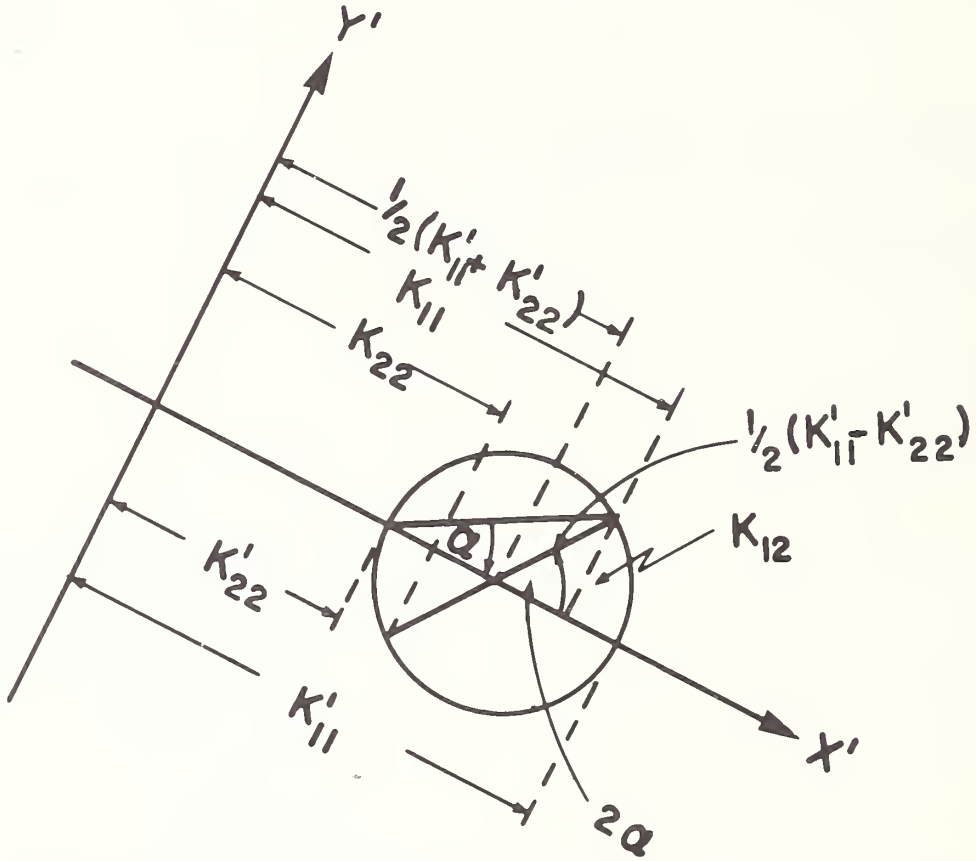

Figure 6.6.--Diagram illustrating details of transformation of axes for anisotropic soil.

The differential equation can be found by applying the principle of conservation to equations 6.20 to 6.23. Having the width of surface stream tube b (chapter 2) the flux in the S direction is bq_s per unit thickness of the profile and the normal flux is accordingly bq_y. Using c for the moisture content, the differential equation is then

$$\frac{\partial}{\partial S}\left[bK'_{11} \frac{\partial}{\partial S} (z + \pi)\right] + \frac{\partial}{\partial y} \left[bK'_{22} \frac{\partial}{\partial y} (z + \pi)\right] = \frac{\partial c}{\partial t} . \tag{6.27}$$

As $\partial z/\partial S = -\sin\alpha$ and $\frac{\partial z}{\partial y} = \cos\alpha$ we get

$$\frac{\partial}{\partial S}\left[bK'_{11}\left(-\sin\alpha + \frac{\partial\pi}{\partial S}\right)\right] + \frac{\partial}{\partial y}\left[bK'_{22}\left(\cos\alpha + \frac{\partial\pi}{\partial y}\right)\right] = \frac{\partial c}{\partial t}. \tag{6.28}$$

We may assume that b and cos α do not change with y. The very essence of our treatment is that we may consider the profile as a whole. Thus, we may wish arbitrarily to assume that K'_{22} does not change with y either. Different assumptions of this kind may lead to different models. They can be interpreted as different types of averaging according to different measuring possibilities.

Using the (zx) system, we obtain the differential equation

$$\frac{\partial}{\partial x}\left[bK_{11}\frac{\partial\pi}{\partial x} + bK_{12}\left(1 + \frac{\partial\pi}{\partial z}\right)\right] + $$
$$+ \frac{\partial}{\partial z}\left[bK_{21}\frac{\partial\pi}{\partial x} + bK_{22}\left(1 + \frac{\partial\pi}{\partial z}\right)\right] = \frac{\partial c}{\partial t} \tag{6.29}$$

6.8 A Preliminary Estimate of Horizontal Conductivities

Strictly speaking the detailed solution of flow in the profile is a two or three dimensional problem and a nonsteady one. However for a first estimate of the horizontal conductivity, one may assume a certain vertical flow regime and estimate K distribution in the profile. This may serve as an estimate of the parallel transmissivity. A better approximation may be obtained by actually calculating the parallel components of flow in the profile from an assumed solution of vertical flow. If the integrated parallel flow is Q_p across a profile D, the anisotropy can be calculated by using the simplified equation

$$Q_p = \int U q_y \tan\alpha \, dy \tag{6.30}$$

or

$$\bar{U} = \frac{Q_p}{\int q_y \tan\alpha \, dy} = \frac{Q_p}{D}\frac{1}{\bar{q}_y} \tag{6.31}$$

where

$$\bar{q}_y = \frac{\int q_y \tan\alpha \, dy}{D}. \tag{6.32}$$

Clearly, $\bar{U}$ will be a function of the q_y distribution in the profile.

One can measure the q_z and π distribution in the profile in an actual field experiment, assuming that there exist vertical partitions in the z direction, the horizontal gradients at the moment these partitions are taken off would be $\partial\pi/\partial x$.

$$\frac{\partial\pi}{\partial x} = \frac{\partial\pi}{\partial z}\frac{\partial z}{\partial x} = -\tan\alpha\frac{\partial\pi}{\partial z}. \tag{6.33}$$

If $\partial\pi/\partial z$ is negative (a permeable layer above a less permeable one) and tan α is negative (z of the interfaces reduce with x), $\partial\pi/\partial x$ is negative and flow is downhill. The total horizontal flow is then approximately

$$Q_x = \int K \tan\alpha \frac{\partial\pi}{\partial z}\, dz. \tag{6.34}$$

We equate Q_x arbitrarily to a flow due to a simplified model

$$Q_x = \frac{(U - 1)\tan\alpha}{1 + U\tan^2\alpha} \int q_z dz. \tag{6.35}$$

Thus, the equivalent U can be found for every flow condition which is actually measured by some infiltration experiment.

For some steady state problems one can obtain this figure by analytic methods. If $K(\pi)$ is given by a formula or as a series of numbers, the vertical flow may be solved by a simple integration (4),

$$z_2 - z_1 = -\int_{\pi_1}^{\pi_2} \frac{d\pi}{q_z/K(\pi) + 1} \tag{6.36}$$

$$\frac{\partial\pi}{\partial z} = - q/K(\pi) + 1. \tag{6.37}$$

Thus

$$Q_x = -\tan\alpha \int_{z_1}^{z_2} \left[q_z + K(\pi) \right] dz = -\tan\alpha \left[q_z D + \int_{z_1}^{z_2} K(\pi) dz \right]. \tag{6.38}$$

From equation 6.35,

$$Q_x = \frac{(U - 1)\tan\alpha}{1 + U\tan^2\alpha} \bar{q}_z D. \tag{6.39}$$

From equations 6.38 and 6.39, one can solve for U:

$$\bar{q}_z D \left[\frac{(U - 1)}{1 + U\tan^2\alpha} + 1 \right] = -\int_{z_1}^{z_2} K(\pi) dz. \tag{6.40}$$

If we have a nonsteady flow problem with varying rain, we can still obtain some idea of the anisotropic behavior. For steady rain the wetting front progresses and increases the thickness D and increases the right-hand side integral with time. The chances are, however, that U will not change significantly, although the total horizontal flux Q_x will increase in proportion to the

depth D. If the rain intensity increases, q_z increases accordingly and so does K (π).

As a first guess, near the surface one can use the rain itself as the vertical flux. The horizontal flux is proportional to the vertical. Thus, the error associated with estimating q_x is the same percentage as the error in the precipitation measurement. Further, as the horizontal flow component is often but a small fraction of the vertical, the error in determining U in this manner should be small.

EROSION

7.1 Introduction

Man's involvement in problems of soil erosion is as old as, or even older than, the terraces found on the Judean mountains in Israel. Even today, the practices and methods available for coping with erosion surpass the actual understanding of the process. These practices have been developed mostly by trial and error. While we have struggled to obtain a better understanding of the physical phenomena involved, some of the incentives to study it have disappeared. Thus, the major breakthrough in the field of erosion is that we have become less dependent on its prevention.

Erosion control and reversal were used to maintain thick enough soil on steep lands. The modern agricultural technology and the modern standard of living often make it uneconomical to cultivate such sloping land. Erosion and sedimentation used to endanger the fertility of the soil by carrying away or covering up the topsoil. Modern fertilization has taken the edge off this problem. Former unproductive land is being turned into cultivated land.

Still, the general field of erosion and of soil structure is of utmost importance for many old and new reasons. Control of streams, pollution, protection of roads and hydraulic structures, and preservation of landscape have become both necessary and feasible for a wealthy society. In view of this, we may have to reexamine the order of priorities in these problems. With a growing population and limited resources, some form of soil conservation or soil reproduction may become again a necessity.

The mechanisms involved in the erosion process have not changed. Before a physical model of erosion is suggested, these mechanisms should be discussed, because the lack of clear understanding of these fundamentals has been the main obstacle to a more rigorous erosion study. The process of erosion probably involves two main parts.

1. A process of detaching soil fractions from the soil surface.
2. A vehicle to carry the detached particles away.

Considerable erosion is not possible unless both processes are present. Without the first one, erosion will not even start. Without the last one, erosion will be limited and will stop of its own accord by moderating the detaching forces. Erosion processes also can be classified by the media concerned, namely the soil parameters and the hydrologic parameters. Naturally, these two media interact and a complete understanding of erosion is not possible without understanding them.

Finally, one can classify the erosion phenomena by some descriptive parameters as to size, source, location, and shape. Some common terms are sheet erosion of the land surface; gullying; stream channel erosion; mass movement; flood erosion; construction erosion; and mining and industrial waste. Each of these forms is given different names. They are poorly defined and are not distinct in their physical mechanism, tool of observation, or method of control.

The right procedure should be to study first the mechanics of detachment and transportation of soil fractions; then to study the soil structure, topography, and hydrology in order to explain the various forms of erosion and to suggest methods for their control.

In the following, we shall not present a complete picture. We shall, however, attempt a physical account of the soil properties that would prove pertinent to the problem of erosion. Several aspects of flow through porous media

will be discussed, especially seepage forces. These forces seem to be one important detaching mechanism that has been neglected. Only a few practical suggestions will be made to illustrate the feasibility of the physical concepts.

One of the most important subjects to benefit from the study of small watershed hydrology is erosion. As was noted before and as will be shown in the following, the two are closely related. Unfortunately, most efforts in hydrological studies seem directed toward hydrograph analysis which is of little significance for erosion.

7.2 Soil Structure

The term soil "structure" is used in a very general sense. It has to do with the arrangement of soil particles, their grouping into clusters, the strength and stability of such clusters under various conditions, their shapes, and their porosities. For many years scientists have been aware of the importance of soil structure. However, very little has been done in relating soil structure in a physically rigorous way to subjects other than structure itself.

The Soil Survey Manual (56) defines a ped as a naturally occurring aggregate and a clod as an aggregate caused by disturbance of soil with some mechanical means. Fragment is used for a soil portion formed by rupture along a surface of weakness. In the following, we shall use only two terms, an "aggregate" which will be defined specifically and a "fragment" which is any part separated from the soil bulk. In the field description of soil structure, the soil survey manual refers, naturally, to aggregates as having three features: shape, size, and distinctness and durability. As will be shown, the feature distinctness and durability is really made up of two different things, adhesion and cohesion. Adhesion is the tendency for two separate bodies or materials to stick to each other. It infers joining together. Cohesion implies internal strength or forces holding together a certain material from within. It is imperative that these terms be defined in soils in order to characterize their structure.

Considering a small aggregate of clay particles, cohesive forces hold them together from within. The separate aggregates have adhesive forces between them. A single clay crystal is held together by very strong cohesive forces--between atoms. Adhesive forces, often through thin water films, hold the separate particles together.

The distinction between adhesion and cohesion is a matter of scale or point of view. Physically speaking, both cohesion and adhesion are caused by molecular phenomena. Considering soil fragments of a given size, forces within these fragments are cohesive and between them are adhesive. To complete the definition of adhesion and cohesion, one should attach to them a dimension or a possible method for measurement. However, because of the complexity of the problem, different methods may be adopted, each uniquely related to a certain problem.

Definition of soil aggregates.--There are naturally occurring soil fragments of different sizes. The ability to distinguish between such fragments must be based on some average physical entity. This average entity varies in the soil. When it varies so that there is a sharp transition or a maximum or a minimum, it may be used to define a fragment. Usually, we shall not call these fragments aggregates if the varying entity is the color, chemical composition, or even the particle arrangement, although these may be good indicators for the existence of aggregates. The common notion is to attach some

consistency or stability to aggregates, along with a natural tendency to cleave at certain surfaces. Thus, an aggregate is part of the soil that has cohesion stronger than its adhesion. In other words, the aggregate has a tendency to break apart from its neighbors and at the same time to maintain its own form. The term "distinctness" may refer to the first tendency and the term "durability" to the latter. In our terms the "durability" refers to the cohesion and the "distinctness" refers to the difference between cohesion and adhesion of a certain aggregate size. To illustrate, there may be a soil of high cohesion with highly stable fragments, such as in some puddled or remolded soils. Still, if this cohesion is uniform, there will be little or no distinctness of aggregates. Soils with little cohesion or durability may still have distinct aggregates.

Clearly, distinctness and durability are functions of aggregate size, just as cohesion and adhesion are functions of size, both usually decreasing with increasing size. When relatively small aggregates have a cohesion much larger than their adhesion, they are termed "stable aggregates." In fact, certain soil fragments are so strong and stable that under very rigorous treatment--extracting sesquioxides, digesting organic materials, leaching salts, introducing dispersing agents, and vigorous stirring--they do not reduce their size. Such fragments are called "elementary particles" and are thought of as having more or less unique crystallographic structure. A "massive" soil is one that has no distinctness (between cohesion and adhesion) for soil fragments of any size. Such massive soils may have either a high cohesion or a low cohesion; the latter is called single grained. The fragment size of a massive soil is determined by the relation between the disruptive forces and the cohesive energy and is influenced by boundary conditions of the soil mass. By mechanically breaking a massive soil into fragments, one produces surfaces of low adhesion, at least temporarily. Under different conditions, the original cohesion-adhesion-size interrelation may or may not be regained.

The term "good aggregation" is often used loosely. A massive soil is considered as having no aggregates. When slight aggregation forms, it is said that the soil has poor aggregation. Then, there may be a soil with large aggregates and, finally, aggregates may be so large as to be massive again.

Clearly, the aggregation must be characterized by the adhesion and by the cohesion as a function of aggregate size and under various conditions, such as moisture, swelling, exchangeable cations, and salt concentration. By mechanical treatment, the same soil material may acquire different distribution of aggregate sizes and adhesive forces. Therefore, there must be found a way to distinguish between reversible and irreversible fragmentation of soil; that is, the distinction between the existing distribution of adhesive forces and the lost or acquired adhesion through remolding, compaction or other agents.

Many physical properties depend on the soil structure. Only a few will be mentioned here that have some bearing on the subject.

Bulk density.--One may sample naturally occurring soil aggregates and measure their bulk density. In a cracking soil, part of the porosity is distributed in the cracks. Each aggregate size has a given frequency of occurrence and its interfaces contribute a certain part to the porosity. Thus, the bulk density and the bulk degree of shrinking and swelling will depend on aggregate size and the size of the sample.

Changing aggregation.--Aggregation may be changed in several distinct ways: By changing the very nature of cohesive forces; by changing their distribution in space; or by rearranging certain soil fragments so that they

remain distinct or have a tendency to coalesce. Different processes such as freezing and thawing, drying and wetting, mechanical breaking and compacting, and chemical treatments may be evaluated as to their influence on the permanent or temporary distribution of cohesive forces. For example, drying and wetting that creates shrinking and swelling have a tendency to coalesce aggregates in the subsoil but have a tendency to break the soil down to fractions in the topsoil. These tendencies are, of course, correlated with certain aggregate sizes. On the other hand, differential stresses in the subsoil will cause shear strain, followed by reorientation of particles along shear planes. Thus, nonuniformity in the cohesive forces is created in the form of the shiny slick sides which are found in grumosols.

Cracking in swelling and shrinking soil.--A crack occurs when the elongation of the soil exceeds tensile strength. The number of cracks will depend on the gradient of volume change in a direction normal to the soil surface. With high gradients (fast drying), cracks are numerous and small. With small gradients (slow drying, usually in the subsoil), cracks are large and far apart. Surface irregularities and boundaries will influence the cracking pattern. Nonuniformity in cohesive strength will also determine the place where cracking will occur.

Hydraulic conductivity and water retention curves.--Water relations will be greatly influenced by the aggregation. In the context of this report, we should only mention that the meaning of results obtained from soil samples depends on the size of the undisturbed sample as compared with the aggregate size distribution and the pore distribution at their interfaces. One should remember that aggregate size ranges from the elementary particles to the field as a whole.

Erosion - Experience shows that in erosion certain particle sizes are detached from the soil and carried away. The state of aggregation is known to influence the rate of erosion and sedimentation, and a high degree of cohesion seems to be able to prevent any erosion.

7.3 Cohesion

Cohesion and adhesion are the same phenomena and differ only by the scale or point of view. We shall, therefore, refer to soils that may present some internal strength (with no normal stresses) as a cohesive soil. Cohesion in soil mechanics is defined as an initial shear strength at normal stresses extrapolated to zero (50, 51, see also fig. 7.1). The actual procedure involves drawing Mohr's circles from triaxial shearing tests and extrapolating the failure line (after Coulomb) to zero normal stress.

The problem in our case is different. One may wish to know several things. First, what is actually the shear strength at zero normal stress (not by extrapolation)? Second, it is of interest to know the tensile strength T at zero shear stresses. If the internal friction angle is φ, then by extrapolation

$$T = c/\tan\varphi$$

where c is the cohesion found by extrapolation and T extrapolated tensile strength. Again, one may wish to measure T not by extrapolation.

There is a certain energy investment involved in producing new interfaces in a soil mass. This energy is the integral of tensile stress, times the elongation, until the two newly produced surfaces become far from each other.

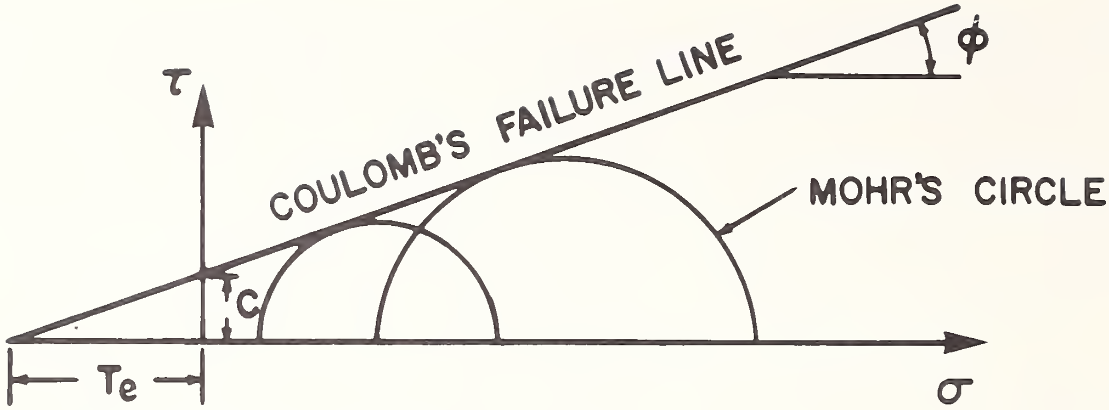

Figure 7.1.--Schematic Mohr's diagram:
σ - normal stresses
c - cohesion, as defined in soil mechanics
T_e - extrapolated tensile strength (force/unit area)

In cohesive soils, this energy may be related to the moisture content and the free energy of water. Evidently, imbibition of water can cause swelling that diminishes the tensile strength to zero (or nearly zero).

A discussion of this aspect is found in chapter 7 of a report by Zaslavsky, Mokady, and Ravina (71). On a pressure plate, drainage can be induced by pressurizing the sample container. The strength of the drained sample is linearly proportional to this draining pressure. When the pressure is relieved but the sample is not allowed to imbibe back the water, the strength is maintained. One can then say that the strength is proportional to the suction or free energy deficiency of the clay water.

It seems that at least in a fine clay the cohesive strength originates in the water films. These films produce a larger cohesion when thinner, or under suction. It is, therefore, easier to detach clay aggregates from the soil when the soil is saturated than when under some suction. In a uniform soil with steady infiltration the suction will be decreased as the rate of infiltration is increased under increased rain intensity (see fig. 6.1). This is an additional mechanism to explain larger erosion by raindrop splashing under larger rain intensity or flooded conditions.

Fragmentation of a cohesive soil by loading.--The failure of ideally elastic bodies under loading can be characterized by the maximum elastic energy that may be stored per unit volume. Disintegration of the material to smaller fragments is governed by the relation between the maximum energy stored E_{max} and the energy E_s necessary to produce a new unit area. If the diameter of the particles produced is D, the area A is proportional to D^2 and the volume to D^3. Then

$$\alpha D^3 E_{max} = D^2 E_s \tag{7.2}$$

where α is a geometric coefficient. Hence

$$D = \frac{1}{\alpha} \frac{E_s}{E_{max}} . \tag{7.3}$$

In the soil, one has to introduce two more aspects. The soil is to some extent plastic and any plastic deformation or a shear causes dissipation of energy. Thus, the rate of loading on the soil is important. The momentarily stored elastic energy will depend on the ratio between the rate of loading and the rate of energy dissipation.

A second aspect which is more important for our erosion studies here is that the cohesion, adhesion, or equivalently derived parameters such as E_{max} or E_s, are functions of the aggregate size.

A rigorous treatment of the elastoplastic behavior of the cohesion, tensile strength, free energy, and water are beyond the scope of this paper. It is important, however, that they be studied rigorously, because the field of application is very large. Trafficability problems, pulverization and aggregation of soil under different speeds, and cultivation at different moisture contents are only some of the problems to benefit.

Some qualitative deductions about erosion.--Different forces tending to detach soil fragments from the soil surface have different functional dependence of the size of the fragment. As will be shown seepage forces are proportional to the fragment volume or D^3. The cohesive stabilizing forces are proportional to the surface area of the fragment or to D^2. Clearly, in this case smaller fragments are more stable. It is easier for seepage forces to detach a large fragment. This deduction is strengthened also by the fact that the average adhesive forces tend to decrease as the aggregate size increases.

Other processes may have other functional relations to the aggregate size. For example, drag forces in flowing water are proportional to D^2 and gravitational forces are proportional to D^3. Thus in sediment transport smaller particles will be more easily carried away.

The two processes mentioned above may produce stable conditions through a natural process. At the beginning, fines will be carried away. Larger aggregates that cannot be transported will produce a natural filter with small seepage forces that will prevent any further erosion. Any change in the hydrologic conditions or through mechanical treatment of the soil may reset the erosion in full motion.

7.4 Seepage Forces and Other Forces Detaching Soil Particles

In the process of erosion, we made a distinction between forces that detach particles from the soil surface and processes that transport these particles away. This distinction is convenient but somewhat artificial. For example, in the movement of adhesionless aggregates as a channel bedload, the two parts of the process seem to become one. The distinction between detaching forces and transporting processes is quite convenient in drawing up strategies of controlling erosion.

The detaching forces include various mechanical stresses that tend to move a particle. We shall mention here seepage forces, shear forces by flowing water, and pounding raindrops. The most commonly mentioned are the shear forces by flowing water. It is probably true that without a certain turbulence particles will not be carried away and erosion will eventually slow down. However, many observations seem to indicate that the water velocity cannot possibly be the primary cause of erosion. For example, eroded rills often start to form where the water velocities are negligible and on very short slopes with very small surface flow. Erosion on river embankments is often

in the form of rills, slumpings, and cave-ins that are normal to the mainstream and where the water is almost stagnant. There are reasons to believe that seepage forces are very important in detaching soil fragments.

Typical drag forces on nonprotruding particles.--Consider a smooth soil surface with some flowing water above it. Roughly, the average shear stress τ between the water and the soil is expressed by

$$\tau = \gamma_W R j \tag{7.4}$$

where γ_W is a unit weight of the water, R the hydraulic radius, and j the hydraulic gradient. To obtain an order of magnitude for τ, let us assume R to be equal to the depth of water in centimeters, and take γ_W = 1g./cm.^{3}. Consider also for this example j, the hydraulic gradient (in a uniform channel), to be 5×10^{-2}. This is a common overall slope of the land and is a very large slope for channels of any significant size. Evidently τ in grams per square centimeters is as follows

$$\tau = 5 \times 10^{-2} R \text{ g./cm.}^2 \tag{7.5}$$

where R is expressed in centimeters. For some reason, a notion is presented that overland sheet flow of water (surface runoff) causes sheet erosion proportional to the shear forces. This proportionality may be a coincidence, if proven to exist, statistically. However, it is hard to believe that it is physically sound. A sheet flow, 1 cm. thick, will produce only 5×10^{-2} g./cm.^{2} shear stress. This stress will not tear off soil with even the slightest cohesive force. Even aggregates without adhesion that are only a part of a millimeter thick will produce enough static friction to resist this shear. Still, a sheet flow 1 cm. thick is an unrealistically huge flood.

Similarly, it may be shown that in most large channels the shear stresses will not exceed a few grams per square centimeter. For example, for a wide channel of 1 percent slope a water depth of 5 meters is necessary to produce a shear stress of only 5 g./cm.^{2}. It is quite common that clay would have initial shear strength (cohesion) of 50 g./cm.^{2} and more.

The situation changes completely as soon as a particle is protruding into the mainstream. Drag forces and moments may be formed that will tear the particle off. This problem has been treated in numerous publications (28, ch. 5).

According to one model, after O. M. White (quoted by Leliavsky, 28), the resultant of the drag force and gravity force becomes critical when it passes through a turning point of the particle. If the drag force becomes larger, the particle will be uprooted or rolled out (fig. 7.2).

As will be shown, the seepage forces can decrease the stabilizing, downward, force in a horizontal soil surface. To maintain the resultant at the turning point, the drag force must be decreased in proportion to the stabilizing force (fig. 7.2).

The seepage forces can be such that the slightest drag would roll out the particle. Adhesion between the particle and its neighbors can produce very stable conditions. The drag forces are proportional to D^2 (D the aggregate diameter). The stabilizing submerged weight is proportional to D^3. Thus, without cohesion, larger (spheroidal) aggregates will be more stable, in proportion to their size.

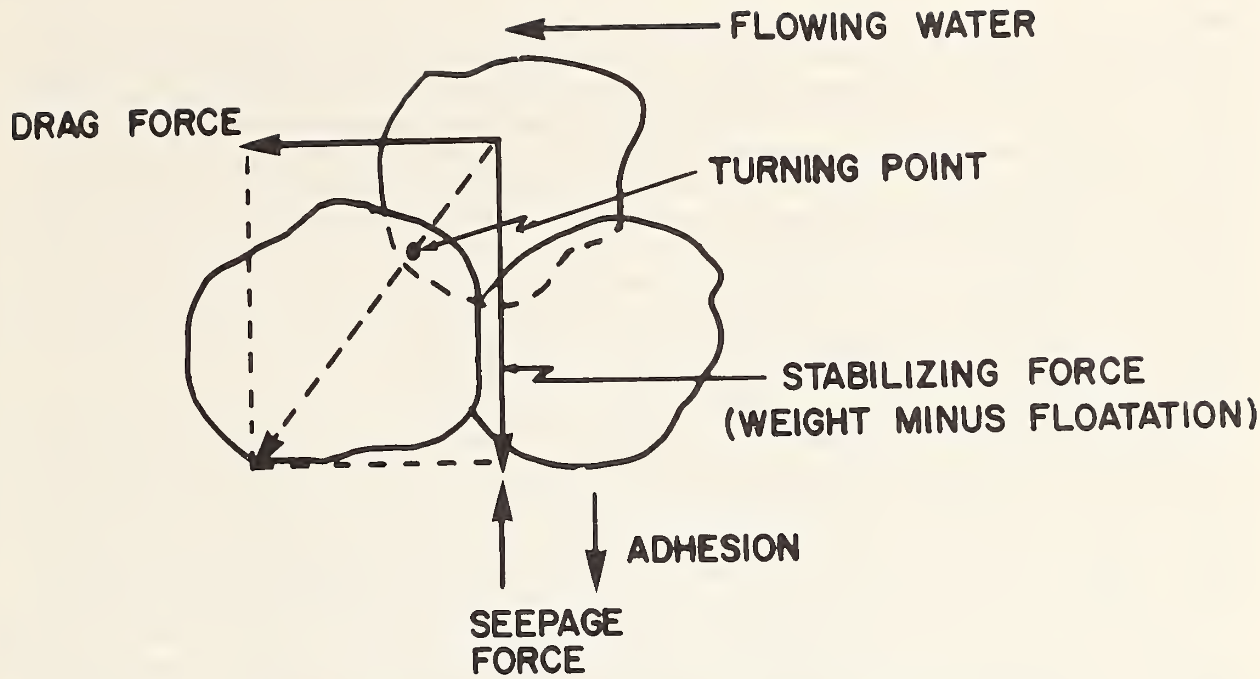

Figure 7.2.--Stability of adhesionless particles against rolling.

Derivation of seepage forces.--Water flow through a porous medium produces shear stress between the fluid phase and the solid phase. The relative motion is at the average rate equal to the flux vector of the liquid. At steady flow, there are equal forces $\underline{F}$ (internal forces) pushing the solid matrix and the fluid phase in opposite directions. $\underline{F}$ is expressed as a force per unit volume of a solid matrix, including pores. The rate of energy dissipation per unit volume is expressed by

$$- \gamma_w \operatorname{grad} \Phi \cdot \underline{q} \tag{7.6}$$

where Φ is the flow potential head and γ_w is the unit weight of water. Equating to the rate of energy dissipation caused by the motion of the solid matrix through the water, we get

$$- \gamma_w \operatorname{grad} \Phi \cdot \underline{q} = \underline{F} \cdot \underline{q} . \tag{7.7}$$

The seepage force per unit volume $\underline{F}$ is then expressed as

$$\underline{F} = - \gamma_w \operatorname{grad} \Phi . \tag{7.8}$$

In an isotropic medium according to Darcy's law,

$$\underline{q} = - K \operatorname{grad} \Phi \tag{7.9}$$

where K is the hydraulic conductivity. Thus, in such a medium

$$\underline{F} = (\gamma_w / K)\, \underline{q} . \tag{7.10}$$

Expressions 7.8 and 7.10 are very important for many phenomena of consolidation, effective stresses (in soil loading), piping, slope stability and erosion. Equation 7.8 was utilized to predict the phenomenon of quicksand (50). When the flow is in upward vertical direction, where the seepage forces F can overcome the submerged weight of the soil, the soil boils up and becomes fluid. We shall later generalize it to nonvertical flow and to soils with cohesion. The use of equation 7.10 can be illustrated in the natural protection against erosion which is produced through elutriation and formation of an inverted filter. First, the fines are carried away. Gravel or large aggregates are left behind so that the hydraulic conductivity K becomes much larger and the seepage forces at the surface reduce. Similarly, a coarse filter protects against piping and is stable against drag forces.

Splashing by raindrops.--The impact of a raindrop can cause a splashing of soil particles. Momentary buildup of the pressure gradients towards the edges of the drop can disintegrate the soil or shoot some particles out. In a way, one can look at this phenomenon as a case of extreme local and momentary seepage forces. However, it is different enough to treat it as a special mechanism. To produce significant erosion, the splashing must throw soil particles into a place where there is a water stream for transportation.

The factors that influence the splashing are (1) the raindrop mass and velocity vector (not only the kinetic energy); and (2) soil properties, such as surface roughness, surface slope and aspect, hydraulic conductivity, moisture content or thickness of water film, particle size-adhesion relation, elasticity and associated mass of the surface. In principle, to solve the problem one has to specify the energy equation, the momentum equation and conservation of mass in an inelastic collision and take into account the capillary phenomena and shock waves within the drop. This is a difficult problem. A dimensional analysis may be necessary in grouping some of the variables with an experimental description of the problem.

Some qualitative predictions can be made. In cohesive soils smaller particles will be harder to detach. On the other hand, to throw out large aggregates, more energy must be transferred from the raindrop. Coarse-grained, dry soil will reduce the momentary buildup of pressure in the raindrop because of surface roughness and faster energy dissipation caused by flow into the soil. Saturation of the soil will weaken the cohesive forces, reduce seepage, and enable stronger splashing. However, a thick enough water layer will protect the soil. Under different conditions different particle sizes will have the highest probability for splashing. Raindrop action can be cumulative, initially breaking aggregates and later throwing them out.

Under many field conditions, distinguishing between the direct influences of the different erosion mechanisms may be difficult. For example, elevating the water table may lead to increased runoff, may produce outcoming seepage in some vulnerable areas, and will make the soil saturated sooner in any given storm and more susceptible to splashing. A furrow or a rill can cause water concentration by shallow flow (see chapters 5 and 6). Water will infiltrate at higher points and seep out at lower ones. Thus, the water will provide at the same time for outward seepage forces, for easier splashing of raindrops, and for a vehicle to carry particles away.

7.5 Stability of Cohesive and Noncohesive Particles

General - To find the stability of a particle near a soil surface three standard tests must be applied.

a. Summation of forces parallel to the soil surface (stability against sliding).
b. Summation of forces normal to soil surface (stability against piping).
c. Summation of moments (stability against rolling out).

The first case is treated in many classical texts on soil mechanics. (For example, 50, ch. 16). It is of special interest to erosion on sloping banks and to form a very loose suspension in a clay soil that is swollen. One may distinguish here between two cases. One is where forces are in the direction of surface water flow and the other is due to land slope and groundwater flow only. The first case is mistakenly considered as the main erosive process. Rather, it is mainly the vehicle of carrying the particles that are already eroded (see section 7.4). Sliding parallel to the soil surface is a special case, valid for long, uniform slopes. It is highly influenced by seepage forces (50). Slumping of banks is a special case which seems to be one of the most important channel erosion mechanisms. In the following, the summation of forces normal to the soil surface will be treated in more detail. The process of detachment in this case is often called piping (70).

Piping of noncohesive material.--Consider a soil surface making an angle α with the horizon or slope $m = \tan \alpha$ (fig. 7.3). Consider also a unit vector $\underline{1}_n$ normal to the soil surface and pointing out of the soil. A flux vector $\underline{q}$ makes an angle ρ with this unit vector. Assuming moderate head differences in the water above the soil surface, equipotentials will be parallel to the soil surface. This will not be the case with a thin layer flowing downhill. The hydraulic gradients will be orthogonal to the soil surface. Thus, any angle ρ between $\underline{q}$ and $\underline{1}_n$ will be different from zero only in anisotropic soil. For the sake of simplicity, only the isotropic and orthogonal case will be treated here. The general case is then straightforward.

The net submerged weight of a particle, for the total volume including pores, F_g is as follows:

$$\underline{F_g} = - (1 - n)\,(\gamma_s - \gamma_w)V\,\underline{1_z} \tag{7.11}$$

where n is the porosity (volume of pores to total volume), γ_s the unit weight of the pore-free solid material, γ_w unit weight of water, V volume of the soil fragment, $\underline{1}_z$ unit vector in an upward direction.

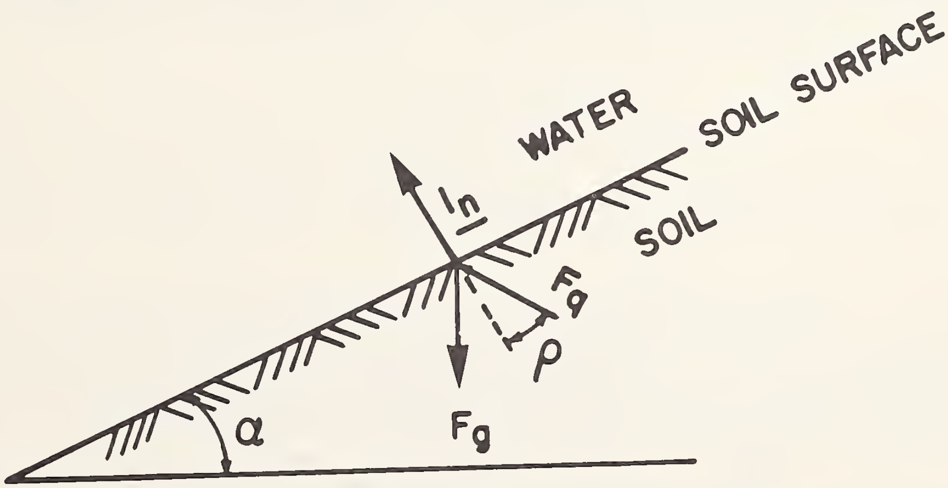

Figure 7.3.--Symbols for derivation of condition for incipient piping.

The component normal to the soil surface (direction of 1_n) is

$$F_{gn} = F_g \cos \alpha = -(1-n)(\gamma_s - \gamma_w) V \cos \alpha . \tag{7.12}$$

The seepage force (assuming orthogonality) is

$$F_s = -V \gamma_w \operatorname{grad} \Phi . \tag{7.13}$$

For flow out of the soil (positive flux q), we have a negative gradient grad Φ and thus, F_s is positive. Combining F_{gn} of equation 7.12 and F_s of equation 7.13 the total active force is

$$F_a = -V\left[\gamma_w \operatorname{grad} \Phi + (1-n)(\gamma_s - \gamma_w) \cos \alpha\right]. \tag{7.14}$$

In a cohesionless soil, this force must be positive to cause piping. In other words, the condition for piping is

$$\frac{-\gamma_w \operatorname{grad} \Phi}{(1-n)(\gamma_s - \gamma_w) \cos \alpha} > 1 . \tag{7.15}$$

This is a generalization of the commonly presented piping formula for a horizontal soil surface (usually called boiling or quicksand). Several conclusions (almost trivial) may be drawn here.

a. An outward flow (grad $\Phi < 0$) may cause piping. Infiltration (grad $\Phi > 0$) is a stabilizing mechanism.
b. The steeper the slope ($\cos \alpha < 1$) the less stable are conditions against piping.
c. Under the same gradients a compacted material (small n) will be more stable.
d. Any mechanism increasing the outward gradient or decreasing the inward gradient will decrease stability against other forces such as drag by flowing water or splashing by raindrops.

In later sections, we shall bring several examples to that effect.

Stability against piping in a cohesive soil.--In equation 7.14 the gravity flotation and seepage forces are summed up. In case there is a net force F_a detaching the particle from its place, there will develop an adhesive force F_c as a reaction. Assuming the maximum tensile stress T between aggregates and a contact area A

$$F_c = a\,T\,A \tag{7.16}$$

where a is some geometric coefficient. For a soil fragment to be unstable the criterion is

$$-V\left[\gamma_w \operatorname{grad} \Phi + (1-n)(\gamma_s - \gamma_w) \cos a\right] > a\,T\,A . \tag{7.17}$$

Rearranging equation 7.17 and putting $(V/aA) = bD$ where D is an equivalent particle and b a geometric coefficient, one gets for instability

$$-\frac{bD}{T}\left[\gamma_w \operatorname{grad} \Phi + (1-n)(\gamma_s - \gamma_w) \cos \alpha\right] > 1 \tag{7.18}$$

To somewhat simplify equation 7.18, we denote an outward gradient by $j = -\operatorname{grad}\Phi$ and state that an aggregate will be unstable if

$$\frac{bD}{T}\left[\gamma_w j - (1 - n)\,(\gamma_s - \gamma_w)\cos\alpha\right] > 1 \tag{7.19}$$

If the outward flux q is known (assuming an isotropic equipotential soil surface), then in place of equation 7.19 one can write

$$\frac{bD}{T}\left[\gamma_w \frac{q}{K} - (1 - n)\,(\gamma_s - \gamma_w)\cos\alpha\right] > 1 \tag{7.20}$$

Many heavy soils may develop tensile strength of up to T - 0.1 kg./cm.² and when compacted even as high as 1 kg./cm.². To make an estimate, one can consider b as being around unity. Using $T = 0.1$ and $b = 1$, we get for instability the condition

$$\frac{bD\gamma_w j}{T} > \frac{bD}{T}\left[\gamma_w j - (1 - n)(\gamma_s - \gamma_w)\cos\alpha\right] > 1 \tag{7.21}$$

or

$$Dj > 100 \text{ cm.} \tag{7.22}$$

It is interesting to note that for $D = 0.1$ cm., j must be of the order of 1000. This is what was actually found in experiments by Zaslavsky and Kasiff (70). They explain why in cohesive soil, splashing by raindrops or swelling and dispersion are necessary to produce appreciable erosion. The momentary outward gradients developed by a raindrop can be very high. Highly dispersed swollen clay has a lower T.

Large soil portions will be more easily detached because of larger D and smaller T. This is really the experience in channels through cohesive soils where large chunks of soil fall out from the bank into the water stream.

In cohesive soils where the water vehicle can carry away only small particles

$$\gamma_w j \gg (1 - n)\,(\gamma_s - \gamma_w)\cos\alpha \tag{7.23}$$

and equation 7.19 can be largely simplified by the following approximation

$$\frac{bD}{T}\gamma_w j > 1 \tag{7.24}$$

which is the condition of instability in a highly cohesive soil.

The momentary gradients produced by splashing raindrops will depend on the permeability of the soil surface. The distance to which a particle will be carried away by splashing will probably depend on D. This, together with equation 7.24, will determine the probability of different size particles being carried away. The process of splashing may, however, be time dependent in the sense that in the beginning a large part of the energy is spent in fragmentation of aggregates that produce smaller aggregates with a low adhesion T.

The essence of soil structure is the distribution of T with particle size D. Thus, soil with a given structure will produce a bias toward a certain particle size. Particles larger than a given size will probably have a much smaller adhesion and thus require an unproportionally smaller j to detach them from the soil surface. Flocculation of clay particles produces small aggregates of very high cohesion but often of a smaller adhesion. Thus, these are probably the smallest particles to be considered as candidates for erosion. Many of the Grumusols have very high cohesion and adhesion to relatively large size aggregates. These aggregates are often too large for a large distance splashing and are sieved out by, or settle in, the small streams where rainwater just starts to accumulate. The erodibility of such soils is expected to be very small. Naturally, a dispersed soil or silty soil with small enough grains of low adhesion are the most erodible by rain.

Stability of particles against rolling out.--To test this process, one has to calculate the sum of moments around a turning point. In this calculation there will be included the moments due to gravity and flotation that are proportional to the volume of the particle. An additional moment is due to drag forces by water, flowing parallel to the soil surface. The drag forces will be especially large on a particle that is protruding into the water stream. From the horizontal and normal forces and the moments, one can find a hypothetical distribution of tensile stresses between a particle and its neighbors. From the earlier section one can find the average tensile stresses normal to the soil surface.

$$\overline{T} \sim bD\left[\gamma_w j - (1 - n)(\gamma_s - \gamma_w)\cos\alpha\right]. \tag{7.25}$$

The presence of the moment can influence the instability in two ways. First, in cohesionless soil, it will roll the particle out as soon as there is a positive net moment (positive moment defined by the direction of rotation caused by the water flowing above the particle or by seepage forces around a turning point).

The moment influences the stability in another way by changing the T distribution between a particle and its neighbors from a uniform one to a triangular one. The actual distribution will depend on the aggregate's shape. Assuming that its projection has a rectangular shape, a rough estimate of the additional tensile stress ΔT is (by simple mechanics) as follows

$$\Delta T \sim \frac{6M}{D^2}, \tag{7.26}$$

where M is the net moment. This is not necessarily the limiting case but can serve as a first estimate. The stress $\overline{T} + \Delta T$ should not exceed the maximum tensile stress. The drag force should be proportional to D^2. As the velocity changes with elevation, a higher power of D is possible. However, the arm of this drag force movement will depend on the protrusion of the particle into the water. Thus, for the same ratio between the protruding part and the part that is imbedded in the soil, the extra stress ΔT will be proportional to D (assuming spheroidal particles).

Seepage forces are proportional to D^3 and the arm of the moment is proportional to D. Thus, the extra ΔT is proportional to D^2. Again, uprooting of particles by rolling them out will depend very strongly on seepage forces and will become more probable for larger particles. A moderate flow into the soil may "suck in" the laminar boundary layer. A flow out of the soil may extend

the thickness of the boundary layer. Thus, it is possible that for small seepage forces the influence on instability will be the reverse of the above. However, in more stable conditions or for high seepage forces, the above prediction will be valid.

There are several theories about the movement of bedload by rolling particles out. In the above, we only introduced two corrections. One involved the passive development of tensile stresses due to adhesion; the other, the introduction of seepage forces. Using any theory on bedload movement, these two corrections should be included.

Both the calculation and observation indicate that it usually takes larger water velocity to roll out a particle than to carry it away once it is out of the soil. This statement must be made with a reservation about certain particle sizes with varying degrees of adhesion. Contrary to this statement, in cohesionless soil with no seepage forces, the smaller particles will be less stable (as drag forces proportional to D^2 and stabilizing forces to D^3). See, for example, a diagram by Schoklitsch (In 26, p. 45). This means that the detaching mechanism in most cases is the rate controlling factor. If the rate controlling factor were the carrying capacity of sediment loads, then all streams would be loaded to capacity, which is rarely the case.

The erosion due to drag forces by themselves or combined with seepage forces and cohesion must be studied in some statistical fashion. On a rough surface there will be some points of weakness where particles will be first uprooted. New surface irregularities will then be formed. The turbulence of the water may produce local and temporal fluctuations of pressure that will render unstable conditions. A rough surface that extends beyond the laminar boundary layer is subject to turbulent fluctuations. There may be formed pressure differences between the upstream and downstream parts of the soil surface irregularities. This will cause less stable conditions in the downstream surface and cause a seemingly backward movement (28, 26, p. 18).

Possibly that a certain success in correlating drag forces and erosion is caused by the correlation between the average drag (proportional to the hydraulic radius R and the head gradient j) and the water velocity. In turn, there will be a relation between the velocity and drag on protruding particles. There will also be a monotonous relation between the velocity times the hydraulic radius and the Reynolds number which characterize the turbulence and boundary layers. So, one can expect some success in correlating average drag forces with erosiveness of a channel. However, to consider the average shear stress produced by the water as the only or immediate cause of erosion would be a mistake. If one looks at the velocity distribution in a channel, the highest drag force would appear at the middle bottom. It is instructive that most of the erosion will start at the bottom side corner where drag forces tend to zero.

7.6 Hydrologic and Topographic Considerations

An internal corner.--An inward corner (fig. 7.4) is a cause for concentration of streamlines and very high hydraulic gradients. A sharp corner forms a cavitation point where theoretically the gradients are infinite. This configuration is quite common in the flow of the ground water into channels. In the absence of a high groundwater table, other mechanisms can produce extremely unstable conditions, such as a perched water table or local saturation followed by a fast drawdown of channel water.

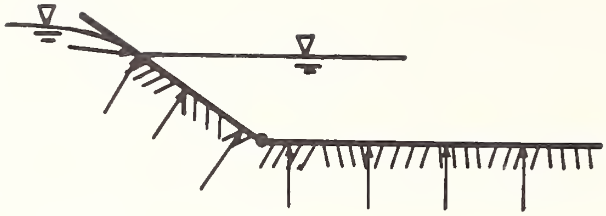

Figure 7.4.--Concentration of streamlines in an internal corner.

A curved boundary and meanders.--A curved soil surface can cause a concentration of streamlines (see chapter 8) and less stable conditions. Even if the seepage forces by themselves cannot detach the soil particles, they increase the probability of detachment by raindrops or rolling out. This case is pertinent also to the formation of meanders in a river.

Natural drains.--A simple calculation can show that for an area of large enough size infiltrating water cannot be carried out only by groundwater flow. There must be some naturally occurring drains with adequate frequency. These drains are in the form of channels, springs, and seepage surfaces. They all involve seepage forces directing out of the soil. A gully is also a naturally occurring drain that may flow only at a high water table following rainstorms or snow melt. The largest streamline concentration will occur at its tip or head. Thus, the chances are that following large storms its head will proceed upstream. This will be the place of largest erosion despite the fact that the water velocities there are often the smallest.

Erosion conditions may develop by high groundwater table or by a shallow impermeable layer. Between the two, the first is more dangerous as it allows a three dimensional concentration of streamlines and no downward leakage. As will be shown, natural drainage into a water pond can produce extremely erosive conditions at the water line.

An embankment on an impermeable layer.--The hodograph method may be used to specify the basic flow characteristics of soil water (4, sections 6.25 - 6.30). With this method, it can be proved that the normal component j_n of the hydraulic gradient causing instability at the bottom of an embankment resting on an impermeable layer (fig. 7.5) is

$$j_n = \frac{\sin(180 - \gamma - \alpha)}{\tan(90 - \gamma)} \tag{7.27}$$

where α is the angle of the impermeable layer with the horizon and γ is the angle between the surface of the embankment and the surface of the impermeable layer. Obviously when $\gamma \rightarrow 90$ $j_n \rightarrow$ infinity. In a steep embankment the stability caused by particle weight becomes negligible. The only protection against erosion in this case can be by introducing a drain in the embankment, by protecting the bottom corner with a coarse and heavy filter or, by combining both. This erosion process can occur in flow layers that are underlain by a heavy impermeable layer. A selective erosion can eventually produce a stable landscape. However, a repeated plowing or reshaping of channels will expose new material to be eroded. A crust very near the soil surface will enable erosion through small grooves in the surface. All these forms serve eventually as natural drains that may be subjected to an increasing rate of erosion (see previous paragraph).

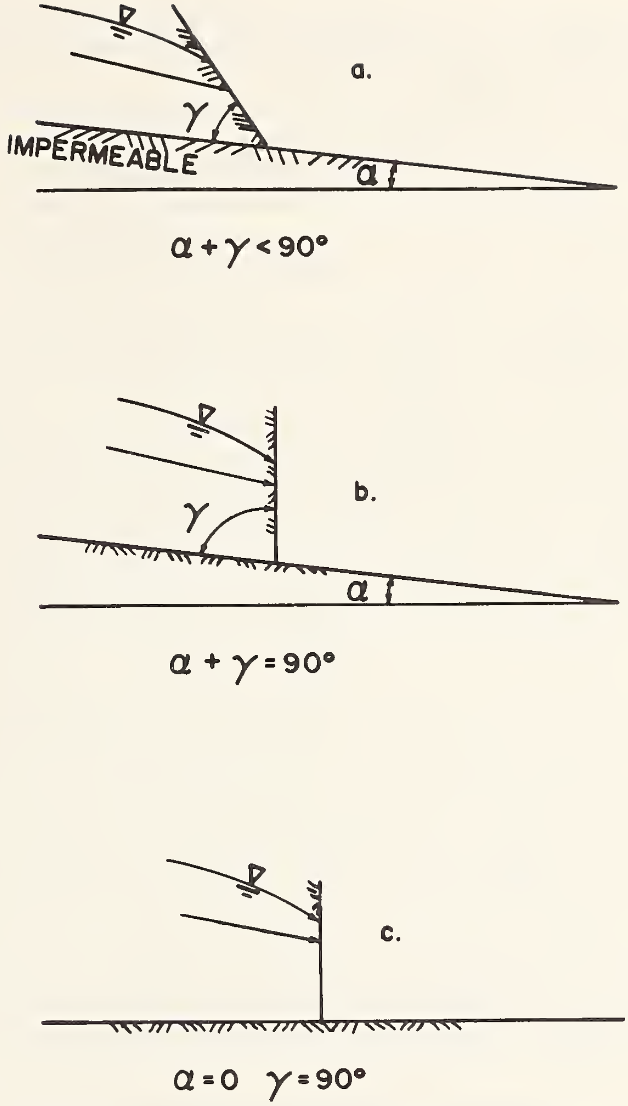

Figure 7.5.--An embankment above an impermeable layer.

Transition between two materials.--Different transitions between soils of different hydraulic conductivities can cause concentration of outcoming streamlines (chapter 6 and 8) and more erosive conditions. In channel flow, a transition between two bed materials is always a dangerous point of erosion.

Ponded water.--By the hodograph method, one can prove that a cavitation point is formed at a water surface where groundwater can seep out. This is a well known experience even where the water in the pond is completely still. Wave action helps, of course, in transporting away eroded particles and in producing pressure fluctuations that increase the instability.

A puddle or a pond is often formed by incoming groundwater. The migration of the water line exposes different portions of the soil surface to erosion. This often may be the reason for the caving in of furrows and the segregation of soil size fraction even in puddles that were not produced by runoff.

Flow through soil layers in parallel.--Water may flow in a more permeable layer downhill. At the bottom it may produce artesian conditions enough to cause unstable conditions.

Flow through soil layers in a series.--In a flow normal to soil layers the largest gradient will occur through the less permeable layer. Thus, a thin membrane or a crust can be more easily eroded whenever back pressure develops behind it. Oppositely, a very permeable layer at the soil surface may produce stable conditions. For a given flux (equation 7.10), the seepage forces will be greater at the surface as K of the surface layered will be smaller.

7.7 Human Intervention in Hydrological Conditions

One may state generally that by raising the groundwater table or by recharging more water into the ground the outgoing seepage forces will be increased. On the other hand, any groundwater pumping, spring interception or underground drainage will reduce this danger and will add to the hydrological stability against erosion. The above is, of course, an oversimplification. Where changes in the moisture regime may go to the dry side, some changes in vegetation may follow that in turn will intensify erosive processes.

In the United States most irrigation, domestic, and industrial water use comes from surface water. The overall effect is probably of more water retention and temporary storage in the ground. To correlate such changes in amount and location of seepage with accelerated erosion by gullying would be very interesting. The operation of water retention, irrigation, and recharge projects will increase runoff at some parts of the flow system or during certain rainy periods. We can conclude that the state of groundwater table and underground drainage conditions are parameters to be considered for an erosion equation or treatment against erosion (62). Models and simulators must also include this aspect.

Change in the soil moisture regime because of different agrotechnical practices will also have a direct effect on the soil parameters and an indirect effect through the changes in the moisture that precedes the rainstorms.

7.8 Some Topographic Aspects

Kirkby and Chorley (22) gave the most intelligent deviation from the usual parameters of average slope and slope length. They emphasized the interrelation between the topography, drainage scheme, and the eventual moisture distribution along the slope. Two quotations from their works are cited below.

"Distance of overland flow is therefore a much less important variable than Horton assumed in determining the zones of maximum flow on many hillsides, and hence the areas most liable to erosion by running water. Under conditions of thick soil cover four other variables seem to assume much greater importance in their influence over erosion.

1. Contour cavities - Producing convergence or divergence of flow (surface or subsurface). It is thus essential to consider three-dimensional slope forms rather than two-dimensional slope profiles.

2. Variations in gradient - These have effects similar to (1) flow tending to be greatest in profile concavities.

3. Variation in soil thickness - These influence the total waterholding capacity of the soil such that patches of thin soil produce saturation and overland flow most readily.

4. Proximity to a perennial stream - In such slope-base locations saturation or near-saturation is common."

In another part: "Overland flow and surface erosion [will occur] at rainfall intensities below the theoretical, general, soil infiltration capacity in the following locations:

1. at the base slopes immediately adjacent to water-filled channels
2. in hollows
3. in slope-profile concavities
4. in areas of thin or less permeable soils."

This article certainly demonstrates an insight into the picture. Still, one has to introduce a few more dimensions into the problem. The most obvious one is that the runoff is not the only hydrological parameter that determines the rate of erosion. Still, similar topographic features seem to influence erosion and seepage forces. Thus, most conclusions will remain unchanged. One should also generalize the picture of flow concentration. Our contention was that flow can concentrate under unsaturated flow conditions as well as under saturated ones. In fact, to distinguish between surface runoff and shallow flow in the soil is physically impossible and analytically unsound. Once the areal unit under consideration is defined, any flow within the area must be taken as one shallow flow. The soil is then a three dimensional medium which is anisotropic.

With a very small furrow or gully, rainwater will concentrate towards the invert before infiltration capacity is exceeded on the average. This concentration will be caused by short bursts of rain of high intensities, by slanting rain and different aspects of the soil slope (in the microrelief), and by anisotropy of the soil. Most of the conclusions reached by Kirkby and Chorley can be drawn from the equation presented in chapter 6 of this paper by observing the sign of the vertical flux in a converging flow system.

At the inverts of the soil microrelief there will be high water concentration. Before long there will appear a saturated state at the bottom of the slope and runoff while above it there may be an unsaturated state. The point is that this process is fast, and the delay time is small, when the microrelief is considered. In many flood hydrographs, such a mechanism is needed to physically explain the rates and volume of runoff. The consideration of the overall slope is not sufficient. In fact, trickles of water streams can often be seen to form in the field with no such thing as a film overland flow and long before the rain exceeds the infiltration capacity.

The above model is also necessary to explain the importance of erosion by splashing. Only this mechanism in the microrelief can provide for a frequent enough vehicle to concentrate and carry away the large amount of splashed material. One can conclude that the microrelief is very important in the hydrology-soil-erosion interrelation. This is undebatable. There seems now to be a way to approach the problem in a physically sound way.

Chapter 8 ANISOTROPY AS A POSSIBLE FACTOR IN SOIL PROFILE DEVELOPMENT

8.1 Introduction

In soil genesis and classification literature, five major factors are considered responsible for producing a wide variety of soil types (18). Following these ideas a more detailed study of horizon differentiation in a profile was undertaken (47). According to Jenny (18, 19) and Arnold (3) a unique combination of parent material, living organisms, topography, climate, and time will tend to produce a unique soil type. Although this seems to be a reasonable tentative statement, its implications should be examined more closely. For a functional relationship between the soil and the soil forming factors to exist, the latter must be well-defined, independent, and sufficiently comprehensive. At the same time, the soil forming factors should be measurable directly rather than by inference, so that the functions can be expressed as a total differential. Given some specific soil type with a property P and some soil forming factors x_i,

$$dP = \sum_i \frac{\partial P}{\partial x_i} dx_i \quad . \tag{8.1}$$

For a given soil type, there likely will be a number of different properties P_j and equation 1 will have a matrix as follows:

$$dP_j = \sum_i \frac{\partial P_j}{\partial x_i} dx_i \quad . \tag{8.2}$$

If the matrix $(\partial P_j / \partial x_i)$ has a non-zero determinant, we may consider the x_i's independent. This matrix may indicate a possible interchange of measurable and deduced parameters by various transformation. It is practically impossible to account quantitatively for all the terms. Still, the preservation of the logical consistency which is described by the above matrix is important.

The five soil forming processes as classically defined do not satisfy the above requirement. For example, it is seldom clear, whether time should be taken as a chronological measure (as one of the x_i) or merely as a descriptive expression of the degree of soil development (as one of P_j). If it is taken as the former, $t = 0$ can be arbitrarily decided upon. Both parent material and relief would constitute the initial conditions that are unlikely to be defined or measured independently back at the time $t = 0$. If time denotes only the degree of soil development, then time is measurable in terms of soil properties P_j themselves and cannot be considered as an independent parameter. It is not possible to give any initial conditions from which time is measured and landscape evolves, since the only way to determine relief and parent material at $t = 0$ is to infer something about them from the present state of the soil. Soil formation is, to a great extent, an irreversible process. Many different initial conditions may lead to similar end results, and it is not possible to project back uniquely from present to past. Similarly, neither climate nor living organisms can be considered independent since they interact with both the parent material and relief.

Against the background of this complexity, it will be attempted to single out the process of water infiltration. Intuitively, it is believed that it affects the differentiation of soil types in a catena along a slope, derived from the same parent material. Because of this limitation, any attempt to determine the total effect of infiltration directly will be very questionable. However, it is reasonable to assume that infiltration implements soil development in some way. To make a differential comparison between soils that have different infiltration rates is then not unreasonable. Moreover, infiltration is a mechanism with which we can experiment. This may result in an oversimplification of the soil forming process. The significance of such formulation may lie, however, in its application to contemporary hydrologic problems.

Certain aspects of the role of water in soil formation were considered by Marbut (33), Norton and Smith (39), Ellis (11), Mattson and Lonnemark (34), Jenny (18), Jaffee (17), Lotspeich and Smith (30), Thorp and coworkers (53), Stabbe and Weight (49), McCaleb (35), in the 7th approximation, supplement (48), Al-Janabi and Drew (2), and many others. In general, it is observed that the differentiation of the horizons is more extreme and that the B-horizon is thicker on flat land or at the foot of a slope than on sloping land or at the crest of a slope (1, 9, 13). Similarly, the degree of soil genetic development, considering the differentiation and thickness of B-horizons, is more pronounced in humid than in arid zones. Contrast, for example, the degree of development found in desert-type soils with that found on Podzols (17, 47).

It is generally conjectured that runoff increases in a downhill direction depriving upper slopes of some potential infiltration and contributing more to the lower ones (11, 18). Conceivably, runoff may also cause erosion and expose successive layers of soil before any significant degree of development in place is achieved (37).

Closer examination of the above concepts raises several difficulties. In many soils the main layer that impedes infiltration and that is instrumental in producing runoff is the B-horizon. However, this horizon is supposed to be a result of the profile development. The cause for differentiation should have been there before the B-horizon was fully developed. Virgin parent material or parts of slope with limited B-horizon development are often very permeable. Under natural conditions, runoff may constitute only a small fraction of the rain--in most cases less than 10 percent. If varying amounts of infiltration and runoff in the traditional sense were the main cause of soil differentiation, then only 10 percent of the rain or less would need to account for it. However, when regions of varying rain intensity are compared, it is found that a wider range of precipitation is required to explain differences in soil formation. Often, similar soils occur in areas of varying rainfall (see for example table 1 in Sherman and Alexander (46).

If erosion were limiting profile development on slopes, then a distinct B-horizon should not be found on sloping land. Although, in some cases this happens to be so, usually one finds as one goes upslope a gradually thinning, but distinct, B-horizon at a given depth (see, for example, Norton and Smith (39). Only subsequent erosion, after the B-horizon had developed, could explain cases where it is altogether missing or exposed at the surface.

A greater difficulty appears when a B-horizon is observed at the lower part of a steep slope which is thicker than the B-horizon found on the nearly

level upland (1, 30). On the lower part of a steep slope, the runoff should be larger and erosion more intensive (16, 22). According to the common views, the B-horizon should have been the thinnest there.

Observation of runoff in the field indicates that water does not flow over the soil surface in the form of a film for any great distance. Very soon it concentrates in small rills and gullies. Once there, the water comes off the land relatively fast. If it ever infiltrates the soil, it does so only in small amounts at highly localized beds or on flood plains where alluvial soils are formed.

An example of B-horizon development as a function of infiltration are the sand dune soils in Israel (20). Because the initial rate of infiltration is high, no runoff in its usual sense is at all possible. Considering infiltration to occur vertically downward, it appears that every part of a sand dune would receive exactly the same amount of water. Since no clay is originally present in the profile, there are strong indications that the clay in the sandy clay B-horizon originates in the windblown dust brought down by rain (21). Therefore, in this case, it may be safely assumed that the rate or profile development will correlate directly with the rate of infiltration. However, on these sand dunes classical catena of soil types can be observed. Young dunes with different degrees of development can also be found (63). Since there was neither runoff nor erosion to start with, a question naturally arises as to how the differentiation of soil types along the slope could have ever begun.

This chapter presents a hydraulic analysis that could resolve some of the above inconsistencies. Several deductions from this analysis should be checked experimentally in the field. If the analysis will be found to be valid both theoretically and experimentally, it may contribute to the quantitative significance of different soil types in watershed hydrology.

8.2 Infiltration into an Anisotropic Soil

Infiltration is usually assumed to be a straight down, vertical, unsaturated flow into the soil. Chapter 6 shows that this is not necessarily the case. The tendency to form a downhill flow has been demonstrated for the case of infiltration in a soil with a more permeable overlying horizon.

In a layered soil, for example, a soil with a well-developed B-horizon, the water infiltrating into the ground will slant downhill. Now, the vertical and horizontal flow components in an anisotropic soil may be calculated.

Consider figure 8.1--if the component of flow normal to the soil layers is q_y and the component parallel to the soil layers is q_s, then their ratio is given by (chapter 6)

$$q_s/q_y = -\frac{U \tan \alpha}{\left(\frac{\partial \Pi}{\partial y} \diagup \cos \alpha\right) + 1} \tag{8.3}$$

where U is the degree of anisotropy, $U = K_{11}' / K_{22}'$. Under steady state, uniform infiltration, or wherever $\partial \Pi / \partial y \longrightarrow 0$, equation (3) reduces to

$$q_s/q_y = - U \tan \alpha . \tag{8.4}$$

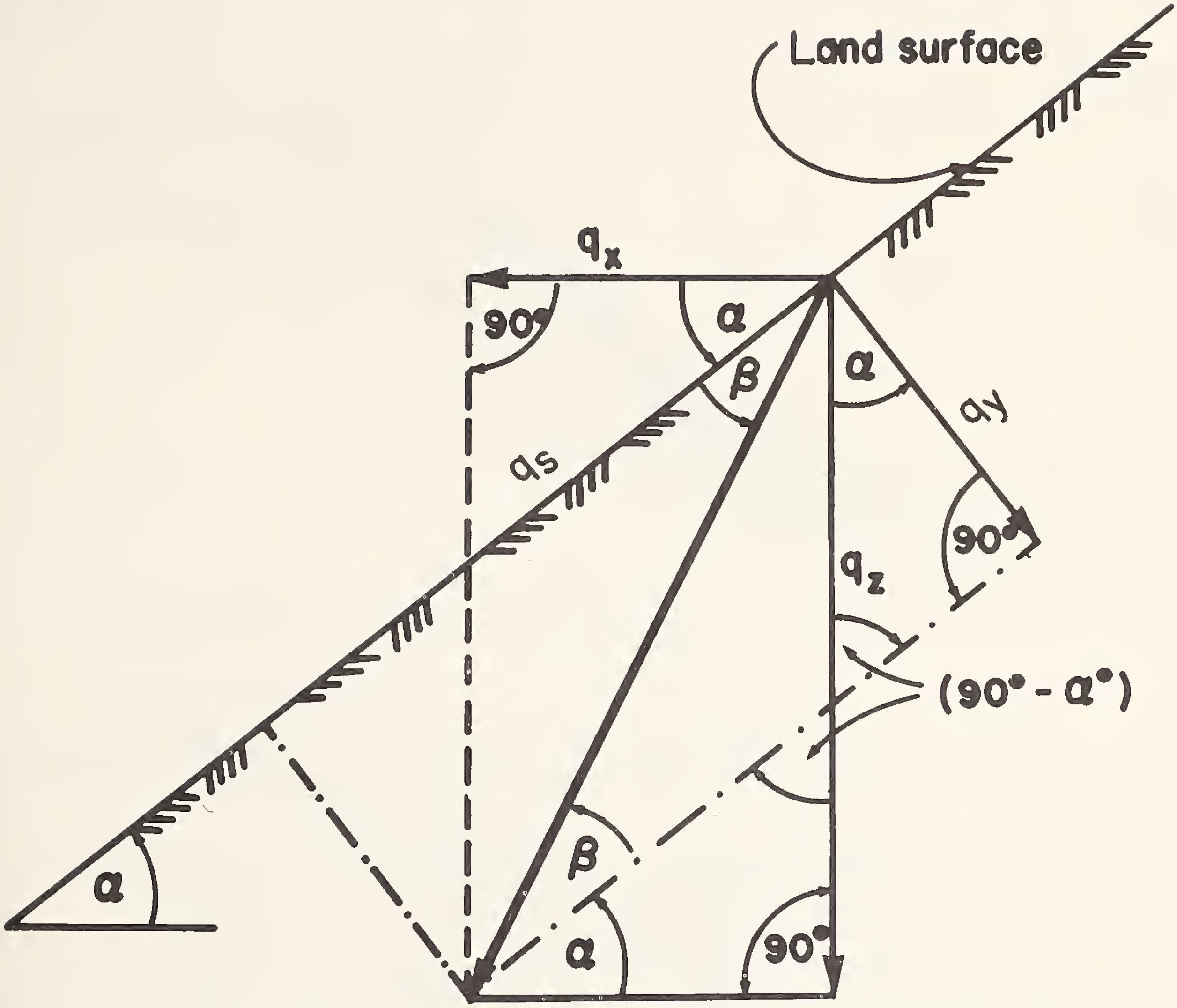

Figure 8.1.--Definition of symbols for infiltration into sloping anisotropic soil.

It was demonstrated that q_s is never zero unless $\alpha = 0$. It should be noted that the angle β between the soil surface and the flux vector can be expressed as

$$\tan \beta = q_y/q_s = -\frac{1}{U \tan \alpha}. \qquad (8.5)$$

Obviously, β decreases as $\tan \alpha$ and U increase. For example, if $\tan \alpha = 0.1$, and $U = 100$, then $\tan \beta = 0.1$. Thus, to reach a depth of one meter in the soil, the water should have to flow downhill for a distance of 10 meters.

Figure 8.2 shows a hypothetical slope. For computational convenience the following equation is used for the slope surface:

$$z = A \tan^{-1} (x/B) \qquad (8.6)$$

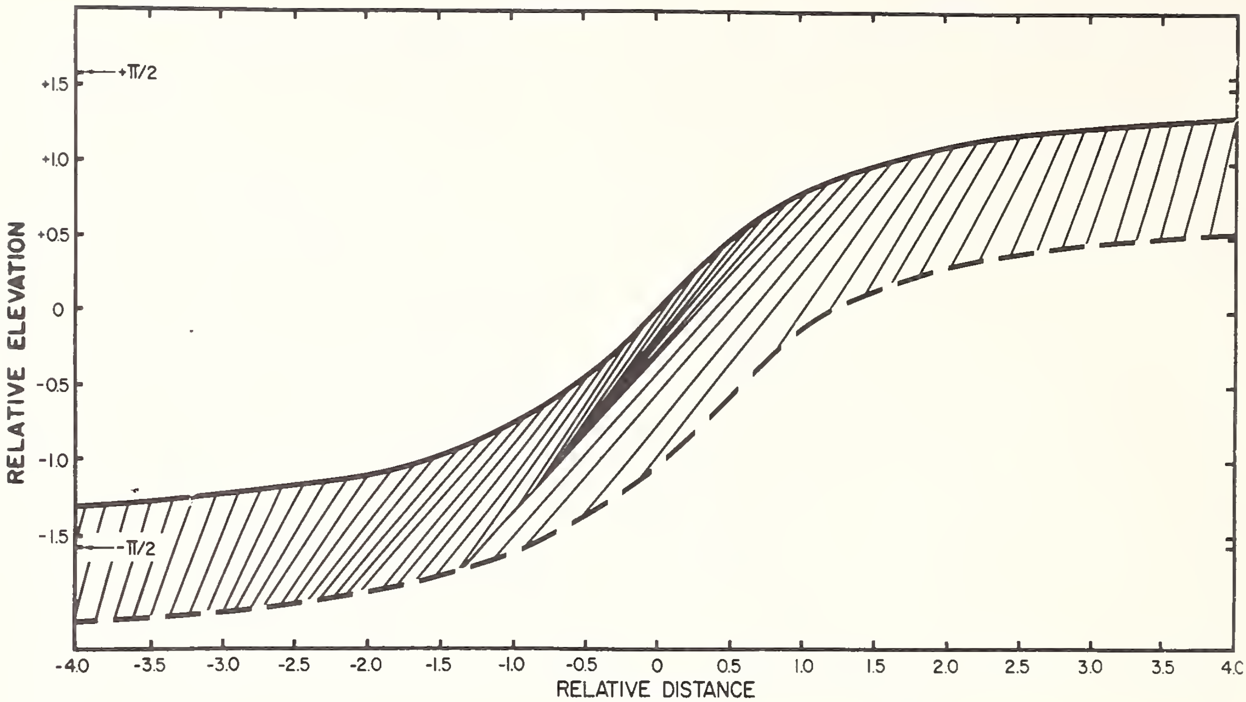

Figure 8.2.--Angle of infiltrating streamlines for slope surface defined by x/B = tan (z/A) and for U = 5.

where A and B are proportionality constants and are set equal to 1 in figure 8.2. The flux directional streamlines are computed from equation 5 for U = 5 figure 8.2.

The use of the constants A and B in figure 8.2 permits the scaling down of any slope of a similar shape. The arctangent cross section illustrates a transition from a convex to a concave slope shape.

In figure 8.3 for U = 10, a relative depth at which streamlines from different parts of the slope converge is shown. Apparently for slopes larger than 1 percent, streamlines will converge at a fairly shallow depth.

In both figures 8.2 and 8.3 the slope is changing. To separate the effect of different U values on a possible direction of streamlines, figure 8.4 was constructed for U = 1, 2, 5, 10, 20, 50, and 100 on a constant 5-percent slope. The implications of the results presented in these figures are further discussed below.

8.3 Consequences of the Anisotropy Theory

Consider some slope where the soil mantle is initially uniform and isotropic with no profile differentiation. Assume that the rain is evenly distributed over the entire landscape and that initially infiltration is the same everywhere. As a result of uniformity and isotropy, the infiltration will proceed straight down (fig. 8.4, U = 1). As soon as profile development is initiated, some degree of anisotropy (U > 1) appears. Assuming that initially U is the same everywhere, the effects of the anisotropy will be more pronounced on the steeper parts of the slope (equation 8.5; fig. 8.2) where the horizontal flow component is larger (β smaller). As a result, on steeper slopes the infiltrating water will be diverted further downhill.

In figure 8.2, 8.3, and 8.4, a dashed line parallel to the soil surface is drawn at some depths. Convergence of the streamlines on some parts of this line and divergence of streamlines on other parts result from interactions

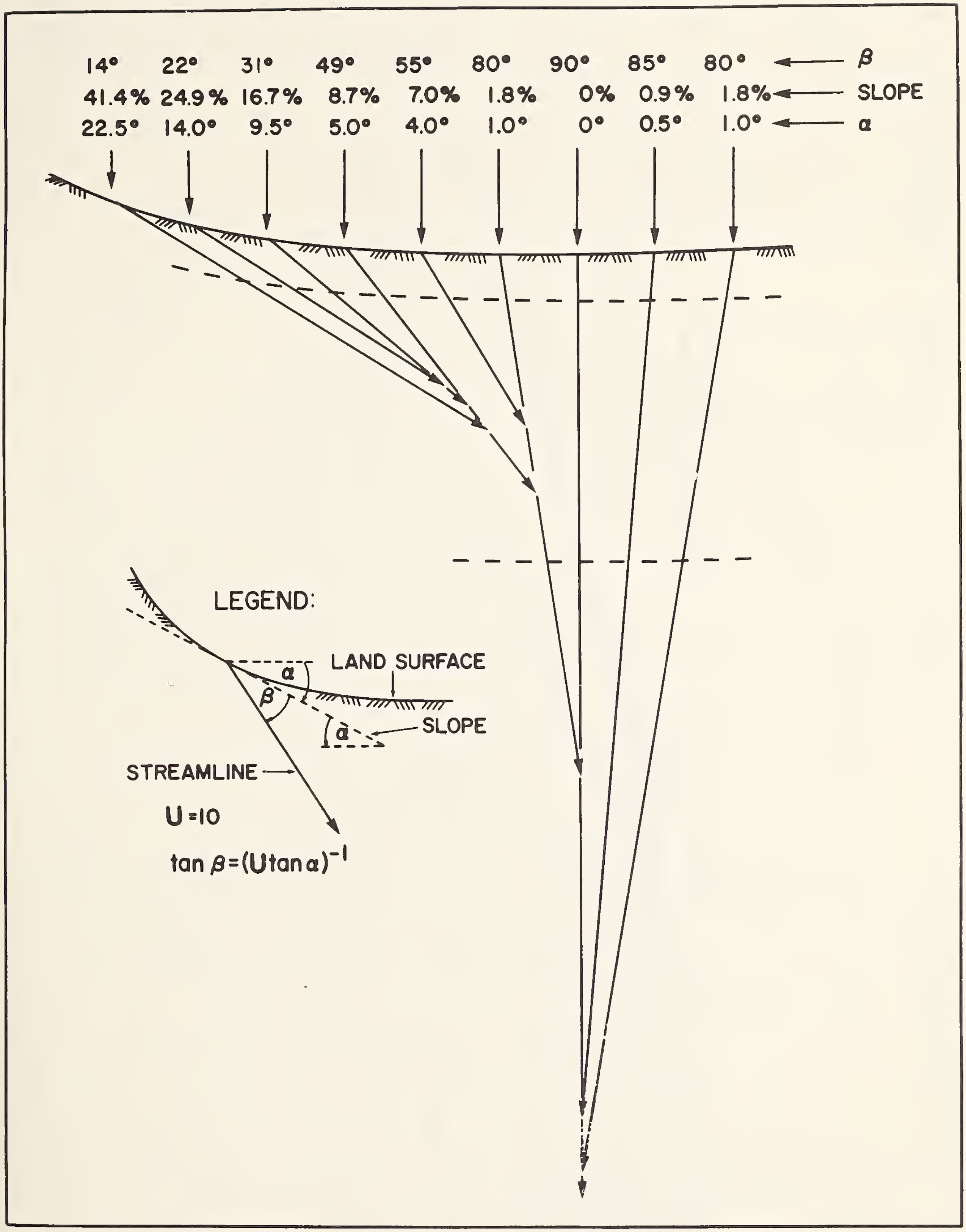

Figure 8.3.--Effect of slope and slope change on depth of layer where convergence becomes pronounced, for U= 10.

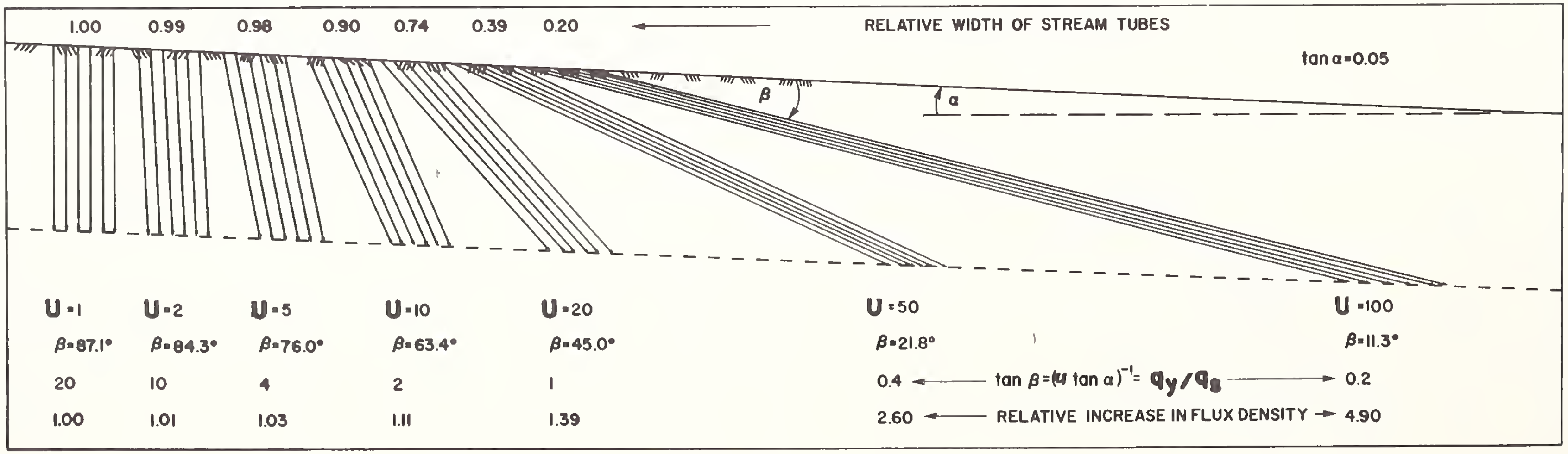

Figure 8.4.--Effect of U on flux direction and density for constant slope.

of anisotropy, slope, and slope changes. Two neighboring streamlines constitute a flowtube. These flowtubes are drawn so that each carries the same amount of water. Thus, the regions where flowtubes are narrower indicate areas that have a higher flux or get more infiltration per unit area (see fig. 8.4). Regions where streamlines are wider apart indicate zones deprived of the infiltration water. The flowtubes narrow down to a point where there is saturated flow or ponded water. For a mixed case of convex-concave slope transition, this can be seen readily in figure 8.2. When slopes are almost flat, the flowtubes converge very slowly to a point quite deep beneath the surface for the particular U used as is evident in figure 8.3. On the other hand, when slopes steepen, shallower zones of flowtube convergence appear as evidenced (see the upper left-hand part of fig. 8.3).

In general, once there is some profile differentiation, the infiltration streamlines will tend to diverge on the convex or more flat soil and to converge on the concave and steep landscape (figs. 8.2 and 8.3). It might, therefore, be expected that the rate of profile development will gradually diminish on a convex relatively flat part of a slope and proceed at a faster rate on a concave part below steeper slopes. This is, of course, only to the extent that infiltrating water is responsible for soil development. This was observed in the field by Young (65) and MacGregor (32) and, in a related topic of certain slope deposits, in the uplands of Scotland by Tivy (54).

Thus, outlined above is a process that can be responsible in part for differentiating various soil types in a catena developing on the parent material. Eventually, a "mature" landscape will have a soil differentiation and anisotropy developed to the point where any further development by infiltrating water is not likely.

One example of differentiation within a criterion (fig. 8.5) is the Palouse catena of Lotspeich and Smith (30). The shape and thickness of the B_2-horizon

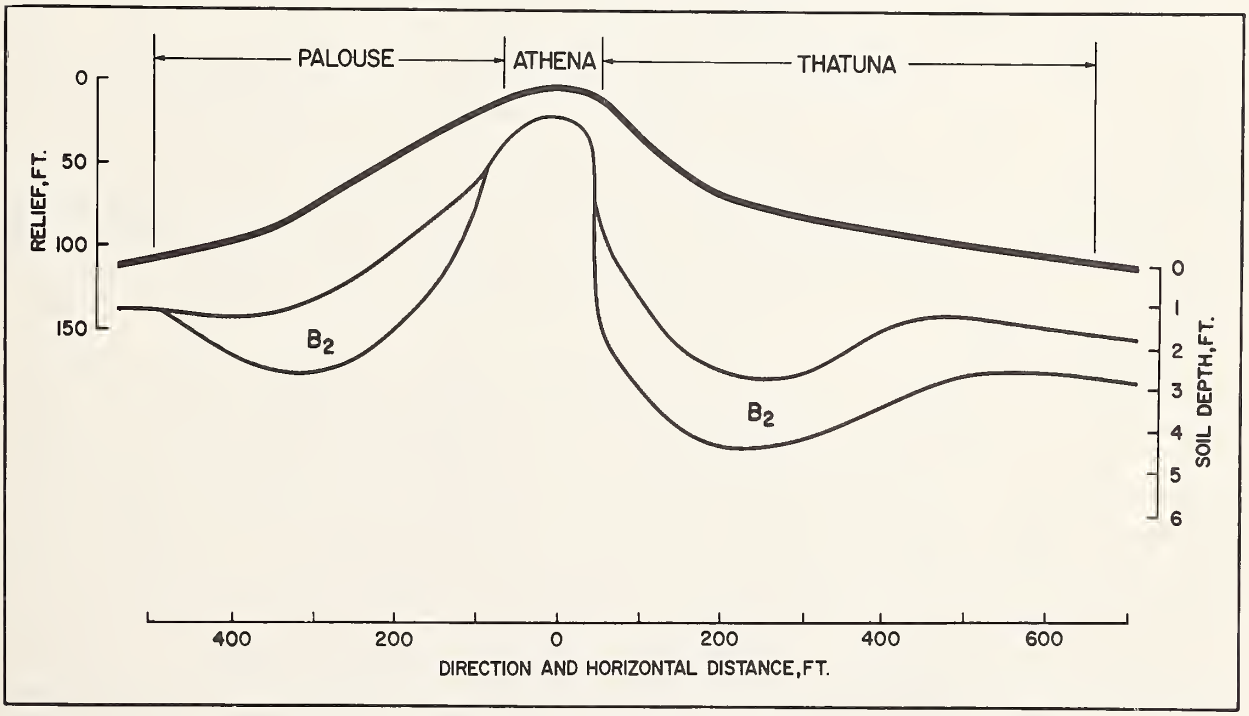

Figure 8.5.--B-horizon in Palouse catena as described by Lotspeich and Smith (30).

tend to support the hypothetical model described here. Another example is the work of Ruhe (42), who considered erosion as a landscape forming process. He observed that the highest and most sharply delineated accumulation of clay was found on the pediment footslope and that soil development and the degree of weathering and poor aeration increased progressively from better aerated slopes to pediment footslopes and finally to poorly aerated floodplain elements of the landscape. The above observations cannot, in our opinion, be explained on the basis of differential pattern of erosion alone. Our model would seem to account for some of the differences observed because of possible leaching patterns, clay migration, and internal drainage within the profile.

Troeh (55) approximated the shape of the land surface with the paraboloids of revolution equations. With these equations he then correlated the soil type and aeration or drainability with the position on the paraboloid. His correlation in order of significance was with the slope, horizontal and vertical curvature of the soil surface. On a hypothetical slope (fig. 8.2) with a given degree of anisotropy, a prediction of drainage characteristics of the soils quite possibly could be made, converging streamlines indicating what would appear to be waterlogged soils or very poorly drained areas.

Our model may have significant hydrological implications. For example, a certain soil type found in a given position on a slope must have an upper limit to its hydraulic anisotropy U. U is the ratio of horizontal to vertical hydraulic conductivity. Anisotropy U is not a constant but is dependent on both the rate of infiltration itself and the water distribution in the profile at a given time. Near saturation, when there is a phreatic surface very near the soil surface, the anisotropy U can become very large. This results in what is usually called surface runoff, because the horizontal hydraulic conductivity at the very soil surface would be many times larger than that of the soil bulk. Although U may not be considered constant, a certain characteristic range of U values should be applicable to each location.

In figures 8.2, 8.3, and 8.4, a two dimensional model is given as an illustration. For proper understanding of the situation in the field, one should draw a horizontal layout of streamlines similar to the approach taken by Aandahl (1). Very often, convergence of streamlines will occur in three dimensions when the topographic contour lines are concave. Wherever streamlines converge, soil profile development is expected to be more significant. The point made here is that this convergence does not necessarily have to occur by an overland flow or even by a saturated flow and that a catena of different soil types can develop of its own accord once some profile development is initiated.

Streamlines generally converge downhill below a concave slope (figs. 8.2 and 8.3). This implies a larger flux rate. As the flow is slanted, the gradients are smaller. Thus, the implication is that a saturated flow must be formed. Moreover, as the hydraulic gradients cannot increase beyond a certain value, positive pressure towards the soil surface may develop. This means that on concave parts of the slope and below them water may seep out and it is more likely there that the rain will turn to overland flow. This can occur even when there are no highly impermeable layers, no perched ground water in the rest of the field, and no high intensity rain that would produce incipient ponding in any other place. Areas where the soil is flooded are more likely to be eroded by raindrops. Runoff will be initiated at such points; moreover, outward seepage of water may, by itself, be an erosive mechanism that would detach particles from the soil mass or make such a detachment by other

mechanisms possible. It is interesting to speculate whether or not the initiation of rills and gullies is related to streamline convergence on steeper slopes and at points of concavities or to some other zones of streamline concentration. Some of these areas are also the zones where a transition to a thicker B horizon would be expected according to our model. A thicker B horizon would tend to impede more downward infiltration and encourage increased surface runoff. The localized development of rills and gullies would provide a drainage system that, in turn, would tend to change the trend in soil development pattern. Downward migration of water would keep concave areas moist and keep bottomland wet for longer periods. This, in turn, would affect the evapotranspiration pattern and concentration of salts in different parts of the profile and could have a marked effect on soil development.

8.4 Other Aspects

Experiments with penetration of fines into sand indicated that most clay particles found in runoff water would stay behind in the first few inches of sand, although some finer particles may migrate to a greater depth. There are probably some geometrical considerations whereby a fine becomes stuck in a narrow pore. A depletion of profile water by plants or evaporation to a point where flocculation of clay occurs should be taken into account. Slow ion exchange in the sand may be responsible for flocculation of clay particles with each other or directly with sand particles. Once some particles are stopped, they tend to stop more effectively the particles following them. In a demonstration conducted at the author's laboratory at the TECHNION, Israel Institute of Technology, a sandy clay from the B horizon of the sand dune catena mentioned previously (20) was treated by leaching with tapwater--almost no clay came out. After being washed with a concentrated NaCl solution, this sample was leached again--this time the leacheate was murky with the released clay. Alternating accumulation of salts and subsequent leaching by rain may possibly be responsible for the clay migration deeper into the soil. According to our model, this would occur not only in flooded bottomlands but also in certain parts of the concave and convex slopes. Carbonates, for example, are considered as inhibitors to the weathering of silicate minerals; once they are removed, the weathering and subsequent migration may proceed. These effects may well be the first step in the differentiation of a profile and may be responsible for a very thick B horizon formed in some concave areas adjacent to bottomland, as distinguished from possible alluvial clay deposits on a flood plain.

The amount of water that infiltrates the soil at any given point should also be taken into account. It will be dependent among other things, on the effective amount of rain available at the soil surface (see chapter 5).

Slopes of varying steepness and aspect will get either more or less rain which would materially influence the amount of infiltration and the degree of development. Furthermore, the rain on a given area may slant near the soil surface away from the horizontal because of the prevailing wind direction. This may cause slopes of a certain aspect to get more rain. There are also other factors connected with the slope aspect such as radiation, evapotranspiration, and cover. Cooper (9), Finney and coworkers (12), and Carmean (1) illustrated some of these; however, these factors are not the subject of the present discussion.

Both cover and landscape microrelief will influence localized concentration of runoff. Surface roughness will produce higher localized rain concentration and will provide faster drainage through higher frequency of concave dips, rills, and gullies. Surface roughness will thus affect the soil differentiation on a macroscale and microscale and reduce the overall rate of soil development. This lack of soil development as a result of faster drainage may be mistaken as the result of erosion. The possibility of landscape denudation by erosion is not excluded here. However, it is shown that runoff and erosion should not be considered as the only mechanisms in the development of a soil catena.

9.1 Infiltration

Infiltration was for many years synonymous with unsaturated flow for some and was considered as saturated by others. The first came from an agricultural background and the last was mainly dealing with hydraulic structures. In fact, both are possible.

In two extremes, the infiltration test by a ring infiltrometer can be shown to be hydraulically questionable. In one case a less permeable layer is developed on the surface which controls the flow (67). In the other case, deeper less permeable layers cause divergence and translation of the wetted body. The test results depend on the size of the area irrigated and on the drainage conditions.

In view of previous discussions, infiltration has to be studied with three dimensional flow in mind. A meaningful infiltration test can be run only with a certain application in mind. In situ measurements with sprinkling and tensiometric measurements may reveal the nature of the soil profile, but there is a lack of practical tools for actual flux measurement. A tracer may be used in some cases. In fact, the direction of flow is frequently sufficient. It renders q_x/q_z, the ratio between components. From continuity, one can determine one of them (for example, q_z) or their sum $(q_x{}^2 + q_z{}^2)$.

An arrangement of pores in the soil causes hysteresis in the pressure-moisture curve. A layered soil can produce another measure of a macroscopic hysteresis (4) both in flow rates and field capacity.

The lack of simple numerical or analytic techniques to solve unsaturated flow problems seems to be a major obstacle in applying flow physics to hydrology. This is true for infiltration, redistribution, and evaporation. A technique for obtaining approximate solutions will be well worthwhile, if it will enable the incorporation of these processes in the engineering practice.

9.2 Evapotranspiration

Evaporation is an important element in the hydrological balance of a landscape. It is necessary to study the definition of this element in order to obtain a well-posed hydrological problem. To illustrate this point, consider first the definition of rain. One can consider a system as a body of soil with some upper boundary surface. Any water falling on this boundary as an input to the system is rain or precipitation. In the same way one may consider evaporation as the amount of water getting out of this surface as a negative input. Several problems are involved here.

The definition of the boundary is not unique, although it must be the same for rain and evaporation. With respect to rain, the boundary is defined by a measuring device, in this case a rain gage. Thus, to be rigorously consistent, exactly the same boundary surface must be chosen for evaporation.

Other control surfaces or boundaries can be chosen. For evaporation this surface is chosen usually somewhere above the vegetative cover of the soil. For a uniform low crop, this poses little problem as the rain gage is usually above the plants. What, however, should be the choice in the case of a forest?

Very often, the rain is measured at the rain gage elevation while evaporation is estimated by measuring depletion of soil moisture. As the evaporation at the two levels is not necessarily the same, an error may be involved. Preventing any evaporation from a rain gage or a snow gage produces results that are not representative of the net precipitation on the soil surface. The source of the problem is that during the rain there is no way of separating evaporation from rain by a direct measurement. Both are inputs into our system although with different signs. Only the net input is measurable or significant.

In practice, it is assumed that during the rain there is no evaporation and rain is a clear evidence of a net input. Accordingly, we may define evaporation as a negative input or depletion of water through a certain check surface above the soil when there is no net input (no precipitation).

However, even this definition fails in some practical cases. Condensation of moisture from the air in a form of dew is very common. Thus, one can obtain a negative evaporation or a net input when there is no clear evidence of precipitation. If a daily average is taken, there will be only a net input evident. The choice of the boundary surface is difficult. For practical purposes, one may choose the soil surface assuming that it will provide for a relatively simple method of measurement. However, in this case one must also perform the rain gaging at this level. Similar problems occur with respect to snow measurement, with consistency in measurements being even more difficult.

An almost unavoidable conclusion is that rain gaging must be done by estimating the amount of water that gets into the soil, and the evaporation would be the amount of moisture leaving the soil. This has, of course, its practical limitation but it is an illustration of the type of rigorous consistency that should be imposed by the hydrologists on soil physicists and climatologists that are responsible for data accumulation.

There seems to be some explanation for the tendency to define precipitation and evaporation with respect to two different boundary conditions and two different scales. The rain is a phenomenon that is too complicated from a physical point of view. About the only way to cope with it is to measure it and consider it as an externally unpredictable imposed parameter. On the contrary, the evaporation is a process that may be traced back to its fundamentals and be predicted through the measurements of some other data. Still, the theoretical considerations of evaporation (and in that sense also dew) must be compatible with the definition of rain with respect to the boundary conditions, initial conditions, and imposed time functions.

There seems to be an urgent need for better cooperation between scientists and engineers who use the same material for different goals or study different factors to be used for the same purpose.

9.3 The Use of Evaporation Data

The knowledge of water depletion from the soil is almost an end in itself, in view of numerous applications. The growth conditions of crops depend greatly on moisture state in the soil. However, there is a growing need to incorporate the dynamics of soil moisture plant relations in some integral parameter. It seems that the ratio between potential evaporation and actual may serve as such an index. Potential evapotranspiration is the "demand" expressing climatological conditions and the actual

evapotranspiration is the "offer" due to the capacity of soil to feed water into the roots. Furthermore, a question naturally arises as to how evaporation data are going to be used by a farmer. Should they be used for policy making, then it is the incidence of drought conditions which should matter. One has to evaluate the "demand-supply" ratio between rainstorms. Is it really necessary to study the physics of evaporation or is it sufficient to have a statistical evaluation of measurements? On an irrigated land the whole effort in terms of policymaking is directed towards estimates of consumptive use of water, and the actual determination of moisture state for a given crop at one farmer's plot proves of less importance. This can be proved, but it is beyond the scope of this report. It would certainly be a useless complication to have measurements of various climatological parameters and soil parameters just to determine routinely a water regime for a given field. In Georgia, where supplementary irrigation may be needed, a decision to irrigate calls for policy decisions and at times an actual determination of the water shortage. However, in southern California or in Idaho where irrigation is regularly necessary, the main problem is not when an irrigation should be initiated, but rather when it should be terminated. It is better to be safe by irrigating more frequently. With fixed irrigation systems and increasing automation, it costs very little water or work to irrigate more frequently. Instead of following the same uncertain average climatologic data or taking moisture measurements, actually observing the wetting front by a simple electrical device would be simpler and more reliable.

The evaporation may be of interest to determine a certain component in the hydrologic cycle, because determining water depletion between rainstorms and the antecedent moisture profile are necessary factors in studying the next rainstorm. However, in this case it is justified to study evaporation only where it is compatible with the procedures used for the study of other hydrologic components. If the rain is taken as a statistical type of input, why should not the runoff be one too? It is then possible to disregard evaporation and infiltration as intermediates. A case can certainly be made for studying infiltration and evaporation if one wishes to make inferences about the state of moisture in the soil as an end to itself. It seems much easier and more useful to study how the water balance may be changed by some changes in evaporation rather than to estimate accurately the absolute values of evaporation.

9.4 Evaporation as One Segment of a Transfer Phenomenon

The process of evaporation consists of a flow in the liquid and vapor phase through part of the soil profile and the air layers near the soil. A well-posed problem in this case must be formulated as follows:

a. Selection of the flow equation, or the relation between certain motivating forces and fluxes.
b. Description of some permanent properties of the medium through which the flow takes place (whether variable conductivity, anisotropy, turbulence etc).
c. Determination of boundary conditions (steady, changing, or fluctuating).
d. Supporting constraints such as equations of state, and conservation of various components of contents, such as water, heat, and salts.

In view of the above steps, we shall point out several consequences that have to be considered. It should be emphasized that while some simplifications can be made in the evaporation equations, none can do away with any of the above. Any simplifying assumption should be made in terms of the above steps. A theory that does not specify such assumptions may be right or wrong but its value is certainly still subject to question.

The flow in the soil is in some parts saturated and in some parts unsaturated. According to Darcy's law, water head is the only driving force and mass flow is the only type of flow. Under many conditions, however, the effects of salt concentration gradients, thermal gradients, and the like should not be neglected.

Similarly, some diffusion processes are as important as mass flow, especially in the vapor phase.

The flow of vapor through the air may be disregarded if the boundary is chosen at the soil surface. However, there are two problems posed there. The first one pertains to the question where the soil surface is and the second is how to specify the boundary conditions at the soil surface. Boundary conditions in this case must be independently measured (vapor pressure, temperature, and maybe fluxes). If they are found by some other theoretical treatment, we come right back to the original problem of considering the soil and the air as one flow system. In specifying the flow equation through the air, one has to know the vapor pressure gradient and the eddy diffusivity because of turbulence at different elevations. To that end, one could try to undertake the study of an extremely difficult problem--the turbulence and boundary layer near the soil. It seems, however, unavoidable to measure the vapor transfer characteristics through some actual evaporation experiment. The big question naturally arises again: Why not measure the evaporation itself?.

Most evaporation and infiltration studies consider one dimensional vertical flow only. As already shown elsewhere in this report, this is not justifiable. While on the average, one may still have a one dimensional problem, it is clear that local variations of evaporation are not only an end to themselves but are also necessary for assessing the effect of a following rain. Variations of evaporation may be on a large and small scale. Nonvertical flow in the soil towards the evaporating surface can be very important and may lead to salt accumulation after a subsequent rain at concave and low points on the surface.

From a macroscopic point of view, the soil will be anisotropic and there will be a very permeable layer at the very surface. However, a layer that is very permeable when unsaturated may become the most impermeable to evaporation.

Many of the boundary conditions at the top boundary are fluctuating with time. There are fluctuations in temperature, vapor pressure, relative humidity, barometric pressure, and certain effects of the wind. One may want to avoid some of these fluctuations by presenting average conditions. This must be done with great care. The values of the potentials and conductivities may be quite different when one averages the fluctuating boundaries and properties of media than for steady or monotonously changing conditions.

The lower boundary is just as problematic. In this case, if ground water is present there will be a movement of phreatic surface. In the case of an unsaturated profile, one could choose a fixed boundary and then have to actually measure the conditions here. Alternatively, it is possible to have an initial condition throughout the profile and bound the medium at the point of

division between the downward and upward flow. This boundary point will be moving with time.

Only under relatively dry conditions with no net recharge from rain, is it possible to consider a limited profile thickness with a negligible interaction with lower soil strata.

The solution of some flow problems in a rigorous way will indicate when and where certain zones can be neglected as to their influence on the solution and when certain forces and fluxes can be disregarded.

A special problem is posed by plant roots. One way is to regard them as distributed sinks. Very probably, these sinks will be a function of the moisture state itself, of climatic conditions and on specific vegetation. The distribution of these sinks is probably not uniform and not steady.

For many problems the influence of plants obviously will have to be simplified. Some limited studies, indicated that lowering of the effective position of the soil surface is perhaps a rough equivalent of the plant influence.

9.5 Pan Evaporation

Very often pan evaporation is used to estimate moisture depletion from soil. A number of difficulties occur in using this type of data. Among them are the nonuniformity of the field and the dissimilarity of evaporation from a small source to that from a large field. By a simple correction coefficient, one can come within 10 to 20 percent of the potential evapotranspiration of the field (in the long run). However, a tremendous effort is needed to improve on that. Moreover, because of nonuniformity of the field, whether a large improvement is meaningful is altogether questionable.

It is easy to show that a reading from a single pan cannot possibly provide for a derivation of actual evaporation from the soil. If, for example, measurements would be made from two pans having two distinctly different vapor pressures one could, at least in principle, separate the atmospheric potential and conductivity. They, in turn, can be used as boundary conditions to the soil surface.

Many practicing agricultural engineers may justly question the need for pan measurements as they are succeeding in determining a figure for an irrigated field, which is rarely in error more than 10 to 20 percent after a season or two or by very rough comparisons. As was mentioned above, the trend in irrigation would be more and more to determine when to close the valve and not when to open it. Many programs of evapotranspiration studies should be critically reviewed in this respect. Certainly at present, there are many data collected with expensive instrumentation that has not proved to fit into any scheme to provide a more useful accuracy in planning or aid in operation or better physical understanding.

9.6 Hydrologic Studies and Pollution

At present, watershed hydrologists do very little in the way of providing data for pollution studies. The closest their data can be applied is by considering the change of total runoff with watershed size. Usually, there is an increase of runoff with size in humid regions, but the reverse seems to be true in arid regions. The change can be from nearly zero to 30 to 40 percent

from a very small surface to a major river, respectively. Any increase is by water that percolates through the soil. The slope of the runoff area curve or the ratio of runoff from a watershed and a subwatershed are indicative of the average path length that different fractions of the runoff make through the soil.

For pollution applications, one has to determine the actual flow net for different flow configurations. Some flow paths of water drops may be miles long and take many years to emerge back to the air. The present soil chemistry studies usually neglect minute or slow processes that may become very significant on this large size and time scale.

Dilution of contaminates can occur in soil by longitudinal and transverse dispersion (4, chapter 10-11. More dilution can occur if a well gets a pollutant only from a certain sector out of its whole periphery.

Stagnation points can be found or produced where a pollutant will stay forever. Pumping of aquifers, drainage, and small dams can be used to change the path length or soil detention time of rain and irrigation water. Places may be defined where pollution can be delayed, reduced by soil processes, or kept to a minimum by proper dilution. All these applications require hydrological studies beyond volumetric statistics or storm hydrographs. Moreover, intentional changes in the surface flow pattern or in ground water regime should prove in many cases a shortcut towards useful information and possible solutions.

9.7 Pumping

In many areas, utilization of ground water by pumping offers a multipurpose solution. Pumping of ground water or shallow drainage can affect different hydrological parameters and can serve also as a tool for research. The main significance, however, of the following discussion is in illustrating an approach to research which is motivated by specific practical problems and narrowed down by available solutions. One can observe induced changes rather than try to measure absolute values. This is a way to obtain a more useful accuracy in our formulation.

Pumping or drainage provides for a local water supply that can often be developed gradually without the need for major investments. Construction of any large surface storage and conduits must be either lagging behind the needs or providing for needs that will materialize only in the future.

In many areas in the United States, the ground water is very near to the soil surface. Drainage or pumping may provide for a larger temporary storage. In some places where there is a semiconfined aquifer, the downward leakage from top layers may be materially increased. This may reduce runoff to the streams as much as is necessary to supplement the water use.

Ground water use can usually provide for large storage that is capable of smoothing out the supply curve over several years. This is often beyond the capability of surface storage.

An added byproduct of ground water utilization is the reduced evapotranspiration from seepage areas through soil or from open water surfaces.

If our views about erosion are correct, then a lowered water table may reduce gullying and other forms of erosion.

To the extent that soil has a purifying capacity, pumping can reduce pollution by forcing larger parts of the flow through longer paths in the soil. Even in the case of pollution by heat, ground water recharge and a pumping

field with proper cooling surfaces provide an excellent solution. While it may not be the most economical for nuclear stations nor quite economical for water supply, the multipurpose combination will be economical.

Even if the soil does not completely eliminate certain pollutants, the soil has a large delay effect that may provide us with a few more decades of a good water supply before new technologies could be applied to alleviate the problem more fundamentally.

The utilization of ground water by pumping is more naturally related to "smaller" watersheds. In view of the above-mentioned aspects pumping becomes strongly related to the intensification of agriculture, water recharge, runoff, pollution, and erosion that are indigenous to small watersheds and agriculture.

9.8 Other Structures for Controlling Ground Water

Surface and underground drains in the proper places may help to stabilize channels and reduce erosion and water losses. In some places, the shallow return flow from rain produces a sizable storage. For utilization of this storage, it is necessary to make the water readily available at a time of need with large enough discharges. For some reason, this problem was always considered in terms of surface storage behind a dam or in a pit. The alternative can be in a long enough drain that can be put into operation when needed. The question of feasibility or economical comparison must, of course, be studied in each case. We must remember that small shallow dams lose much water by evaporation and gain much in salinity.

A permeable rockfill dam that is gradually filled up by silting serves as a plug that slows the return flow by drainage into small tributaries. Such a measure may be detrimental as it will produce some new drainage channels by erosion. However, in combination with wells or drains, such a dam provides for an increased storage or delayed drainage.

9.9 Hydrological Studies for Special Purposes

Throughout this report the writer repeatedly stated that different approaches may be chosen for different ends. Consider, for example, monthly water yields. In many watersheds, one gets a relatively smooth "base flow" that is not influenced very markedly by an individual rainstorm. This base flow may be adjusted by taking into account slight changes in the ground water table or perched water table. If the long range base flow is Q, than ΔQ can be simply estimated in proportion to changes in the ground water table $\Delta\Phi$ (see chapter 5). $\Delta\Phi$ can in turn be estimated by roughly estimating the net recharge of the monthly rain. If ΔQ amounts to 20 percent of Q and if ΔQ can be estimated in the crudest manner to 50 percent accuracy, then ΔQ can be estimated to 10 percent accuracy which is probably 1 to 3 percent of the total rain. Often the immediate storm hydrograph (over and above Q) constitutes a few percentage points of the monthly flow. Thus, even a rough estimate of the immediate hydrograph runoff can still produce a fairly accurate estimate in the monthly flow.

The above is a hypothetical scheme to estimate expected outflow from watersheds. It calls for ground water piezometeric measurements associated with small changes in base flow Q. It becomes obvious that almost any model for monthly runoff estimate will do. If we underestimate the runoff, we shall

get a higher value for ground water that will in the long run compensate for the volume of the total water yield.

More than the numerical accuracy of the runoff hydrograph it is important that the model will be physically sound so that one could use the model for intentional changes in the system. Models have been known to be successful in numerically estimating the monthly yield without the use of ground water measurement. However, they cannot provide for possible changes by irrigation, drainage, pumping, terraces, and so forth.

If one is interested in peak flows or peak yields for short time periods, the above approach will not suffice. In fact the same models that will be successful in estimating the monthly yields make errors up to a few hundred percent in the estimate of peaks.

If the stability of channels is at stake, a fast drawdown of water may be one of the most important parameters. In this case, one is interested mostly in the rate of buildup of a storm hydrograph or in the number of water table fluctuations in the channel flow. More comprehensively, one would probably be interested in the amplitudes and periods of discharge fluctuation (actually their probability distribution) and the coincident distribution of the surrounding water table or seepage from the channel banks.

In this case, maintaining the accuracy of monthly volume estimates is not necessary, but the number of drawdowns and rates of drawdowns is important. The same total water yield may be very erosive in climate with extremely fluctuating discharges (as is in some arid or semiarid zones), but can have little or no erosive power in some perennial streams of more or less steady rates of flow.

Many more examples may be cited on how some specific problems or available solutions can narrow down the experimental efforts in view of the unavoidable need to make approximations of the real world.

9.10 Trial and Error and Followups

A large part of engineering development has been obtained by a gradual narrowing of the necessary tolerances through trial and error. The theory helps us to narrow down the number of unknowns to organize our accumulated knowledge and to teach unexperienced engineers. There is nothing wrong with approximate theories that are not all inclusive.

An extremely large source of knowledge is neglected because of the lack of followup on various projects. It is unfortunate that many of our institutions are compartmentalized according to semantics. In hydrology, the research must be a part of every road building, dam, drainage system, and such. The followup program must be designed as part of the construction. Budget allowances should be made for the followup as a part of the structure. It must be considered as an investment in developing a better product.

Of course, the decision of what and where to follow must depend on the relative importance of the subject, the degree to which it is characteristic to other areas and can serve as a prototype and to the anticipated improvement.

Many examples may be cited. Culverts and bridges can be easily adopted with very little cost for runoff measurements. Underground drainage is the best tool for studying future drainage--hardly any theory is necessary. Soil conservation measures must be followed with measurements of changes in runoff, erosion, or ground water recharge. Records of existing wells should be collected.

REFERENCES CITED

(1) Aandahl, A. R. 1948. The characterization of slope positions and their influence on the total nitrogen content of a few virgin soils of western Iowa. Soil Sci. Soc. Amer. Proc. 13: 449-454.

(2) Al-Janabi, Abdul M., and Drew, James W. 1967. Characterization and genesis of Sharpsburg-Wymore soil sequence in southeastern Nebraska. Soil Sci. Soc. Proc. 31: 239-244.

(3) Arnold, R. W. 1965. Multiple working hypothesis of soil genesis. Soil Sci. Soc. Proc. 29: 717-724.

(4) Bear, J., Zaslavsky, D., and Irmay, S. 1968. The physical principles of percolation and seepage. UNESCO. Arid Zone Res. Ser. 29. Paris.

(5) Betson, R. P. 1964. What is watershed runoff? Jour. Geophys. Res. 69: 1541-1552.

(6) Burgy, R. H. and Luthin, J. N. 1956. A test of the single and double ring types of infiltrometers. Trans. Geophys. Union. 27: 189-192.

(7) Carmean, W. H. 1967. Soil survey refinements for predicting Black Oak site quality in southeastern Ohio. Soil Sci. Soc. Amer. Proc. 31: 805-810.

(8) Chugaev, R. R. 1964. Stability analysis of earth slopes izdatel'stvo "energia." Moskava, Leningrad. Translated from Russian Israel Program for Scientific Translations, Jerusalem, 1966.

(9) Cooper, A. W. 1960. An example of role of microclimate in soil genesis. Soil Sci. 90: 109-120.

(10) Crawford, N. H. and Lindslay, R. K. 1966. Digital simulation in hydrology, Stanford watershed model IV. Dept. of C.E., Tech. Rpt. 39, Stanford University, Stanford, Calif.

(11) Ellis, J. H. 1938. The soils of Manitoba. Manitoba Economic Survey Board, Rpt. 15, Proj. 14, 112 pp. Winnipeg, Manitoba.

(12) Finney, H. R., Holowaychuk, N., and Heddleson, M. R. 1962. The influence of microclimate on the morphology of certain soils of the Allegheny Plateau of Ohio. Soil Sci. Soc. Amer. Proc. 26: 287-292.

(13) Gilbert, G. E. and Wolfe, J. N. 1959. Soil moisture investigation at Neotoma, a forest bioclimatic laboratory in central Ohio. Ohio Jour. Sci. 59: 38-46.

(14) Hewlett, J. D. and Hibbert, A. R. 1963. Moisture energy conditions within a sloping soil mass during drainage. Jour. Geophys. Res. 68: 1081-1087.

(15) Holtan, H. N. 1961. A concept for infiltration estimates in watershed engineering. U.S. Dept. Agr., Agr. Res. Serv. ARS. 41-51, 25 pp.

(16) Horton, R. E. 1945. Erosional development of streams and their drainage basins: Hydrophysical approach to quantitative morphology. Geol. Soc. Amer. Bul. 56: 275-370.

(17) Jaffe, J. S. 1949. Pedology. 2d. Ed., Pedology Publications, New Brunswick, N.J.

(18) Jenny, H. 1941. Factors of soil formation. Ed.; McGraw Hill Book Co., New York.

(19) ________ 1946. Arrangement of soil series and types according to functions of soil forming factors. Soil Sci. 61: 375-391.

(20) Karmeli, D. 1967. A method for sampling heterogeneous soils. Ph.D. Thesis, Hebrew University, Jerusalem (In Hebrew).

(21) ________ and Ravina, I. 1968. A study of the Hamra soil association of Israel. Soil Sci. 105: 209-215.

(22) Kirkby, M. J., and Chorley, R. J. 1967. Throughflow, overland flow and erosion. Bul. the Internatl. Assoc. Sci. Hydrol. 12(3): 5-21. 1967.
(23) Korn, Granino A., and Korn, Theresa M. 1961. Mathematical handbook for scientists and engineers. McGraw-Hill Book Co., New York.
(24) Koshlyakov, N. S., Smirnov, M. M., and Gliner, E. B. 1964. Differential equations of mathematical physics. North-Holland Publishing Co., Amsterdam.
(25) Kraijenhoff van de Leur, D. A. 1966. Recent trends in hydrograph synthesis committee for hydrological research T.N.O. Proc. and Infor. 13 Technical, the Hague.
(26) Landau, L. D., and Lifshitz, E. M. 1958. Statistical physics, Addison - Wessley Series in Advanced Physics.
(27) Langbein, W. B., and Leopold, L. B. 1966. River meanders: Theory of minimum variance. U.S. Geological Survey, Prof. Paper 427H.
(28) Leliavsky, S. 1966. An introduction to fluvial hydraulics. New York, Dover Pub. Reprint from 1969 impressions.
(29) Leopold, L. B., and Langbein, W. B. 1962. The concept of entropy in landscape evolution. U.S. Geological Survey, Prof. Paper 500-A, 20 pp.
(30) Lotspeich, F. B., and Smith, H. W. 1953. Soils of the Palouse loess: I. The Palouse catena. Soil Sci. 76: 467-480.
(31) Lyon, T. L., Buckman, H. O., and Brady, N. C. 1950. The nature and properties of soils. The MacMillan Co., New York.
(32) MacGregor, D. R. 1957. Some observations on the geographical significance of slopes. Geography 42: 167-173.
(33) Marbut, C. F. 1928. A scheme for soil classification. Proc. 1st Internatl. Congress and Soil Sci. 4: 1-31.
(34) Mattson, S., and Lonnemark, S. A. 1939. The pedography of hydrologic Podzol series. I. Ann. Agr. Coll. Sweden 7: 185-227.
(35) McCaleb, S. B. 1959. The genesis of the Red Yellow Podzolic soils. Soil Sci. Soc. Amer. Proc. 23: 164-168.
(36) Meyer, L. Donald. 1965. Mathematical relationship governing soil erosion by water. Jour. Soil and Water Conserv. 20(4): 149-150.
(37) Milne, G. 1936. Normal erosion as a factor in soil profile development. Nature 138: 548-549.
(38) Morse, P. M. and Feshbach, H. 1953. Methods of theoretical physics. McGraw-Hill, New York, N.Y.
(39) Norton, E. A., and Smith, R. S. 1930. The influence of topography on soil profile character. Jour. Amer. Soc. Agron. 22: 251-262.
(40) O'Brien, M. P. 1933. Review of the theory of turbulent flow and its relation to sediment transportation. Amer. Geophys. Union, Hydrol. Sec. Trans., pp. 487-491.
(41) Protz, R., Presant, E. W., and Arnold, R. W. 1968. Establishment of the modal profile and measurement of variability within a soil landform unit. Canad. Jour. Soil Sci. 48: 7-19.
(42) Ruhe, R. V. 1960. Elements of soil landscape. 7th Internatl. Congr. Soil Sci. Trans., (Madison, Wisconsin) v. 23, pp. 165-170.
(43) Scheidegger, A. E. 1967. A thermodynamic analogy for meander systems. Water Res. Res. 3: 1041-1046.
(44) ________ 1967. On the topology of river nets. Water Res. Res. 3: 100-103.
(45) ________ 1968. Horton's law of stream order numbers and a temperature-analog in river nets. Water Res. Res. 4: 167-171.
(46) Sherman, D. G., and Alexander, L. T. 1959. Characteristics and genesis of low humic latosols. Soil Sci. Soc. Amer. Proc. 23: 168-170.

(47) Simonson, R. W. 1959. Outline of a generalized theory of soil genesis. Soil Sci. Soc. Proc. 23: 152-156.

(48) Soil Survey Staff. 1967. Soil Classification, comprehensive system 7th Approximation. 265 pp. Supplement issue, March 1967. 207 pp. U.S. Dept. Agr., Soil Cons. Serv.

(49) Stabbe, P. C. and Weight, J. R. 1959. Modern concepts of the genesis of Podzols. Soil Sci. Soc. Proc. 23: 161-164.

(50) Taylor, D. W. 1948. Fundamentals of soil mechanics. Stability of slopes. 700 pp. Wiley and Sons Inc. London.

(51) Terzaghi, K von. 1943. Theoretical soil mechanics. Wiley, New York.

(52) Thakur, T. R. and Scheidegger, A. E. 1968. A test of the statistical theory of meander formation. Water Resources Res. 4: 317-329.

(53) Thorp, J., Cady, J. G., and Gamble, E. E. 1959. Genesis of Miami silt loam. Soil Sci. Soc. Proc. 23: 156-161.

(54) Tivy, J. 1962. Investigation of certain slope deposits in the Lowther Hills, southern uplands of Scotland. Inst. Brit. Geograph. Trans. and Papers, pp. 59-73.

(55) Troeh, F. R. 1964. Landform parameters correlated to soil drainage. Soil Sci. Soc. Proc. 28: 808-812.

(56) United States Department of Agriculture 1951. Soil survey manual. U.S. Dept. of Agr. Handbook No. 18.

(57) ________ 1957. Hydrology. Supplement A, section 4 of Engineering Handbook. U.S. Dept. Agr., Soil Conserv. Serv. Washington, D.C.

(58) United States Department of the Interior, Office of Water Resources Research 1966. Water resources thesaurus (November) 237 pp.

(59) Vanoni, V. 1965. Review of research activities in sedimentation. In Proc. Federal Inter-Agency Sedimentation Conference, 1963. U.S. Dept. Agr. Miscel. Pub. 970, pp. 8-12.

(60) Chow, Ven Te 1964. Handbook of applied hydrology. McGraw-Hill, New York.

(61) Von Schelling, H. 1951, 1964. Most frequent random walks. Gen. Electric. Co. Rpt. 64 GL92. Schenectady, N.Y.

(62) Wischmeier, W. H., and Smith, D. D. 1965. Predicting rainfall-erosion losses from cropland east of the Rocky Mountains. U.S. Dept. of Agr. Agr. Handb. 282, 47 pp.

(63) Yaalon, D. H., and Dan, J. 1965. Factors controlling soil formation and distribution in the Mediterranean coastal plain of Israel. In Proc. Internatl. Assoc. Quaternary Res. Symp. on Soil Stratigraphy, Nevada.

(64) Yoder, R. E. 1936. A direct method of aggregate analysis and a study of physical motive of erosion losses. Jour. Amer. Soc. Agron. 28: 337-351.

(65) Young, A. 1964. Slope profile analysis. Zeitschr. f. Geomorphologies Suppl. Bd. 5: 7-27.

(66) Zaslavsky, Dan. 1957. Measurement of cracking in a compacted clay soil. M.Sc. Thesis, Technion Israel Institute of technology, Haifa. (Hebrew).

(67) ________ 1964. Theory of unsaturated flow into a nonuniform soil profile. Soil Sci. 97: 400-410.

(68) ________ 1968. The average entities in kinematics and thermodynamics of porous materials. Soil Sci. 106(5): 358-362.

(69) ________ and Buras, N. 1967. Crop yield response to nonuniform application of irrigation water. Amer. Soc. Agr. Eng. Trans. 10(2): 196-198, 200.

(70) Zaslavsky, D. and Kasiff, G. 1965. Theoretical formulation of piping mechanism in cohesive soils. Pub. Geotechnique 15(3): 305-310.

(71) ________ and Mokady, R. S. 1967. Non-uniform distribution of phosphorus fertilizers: An analytical approach. Soil Sci. 104: 1-6.

(72) ________ Mokady, R.S., and Ravina, I. 1966. Development of methods for measuring partial molar free energy of soil water. Annual Report No. 1, Project No. A10-SWC-24. U.S. Dept. Agr. under PL 480. Lab. of Soil Science Technion Israel Inst. of Technology, Haifa. Project Ag.-25.